Structural and functional studies of human ADAM 12 in myoblast fusion and Ebola virus VP40 in assembly

Thesis

Structural and functional studies of human ADAM 12 in myoblast fusion and Ebola virus VP40 in assembly

Joanna TIMMINS

A thesis submitted in partial fulfilment of the requirements of the Open University for the degree of Doctor of Philosophy

April 2002

Sponsoring Establishment: NIMR, Mill Hill, U.K.

Collaborating Establishment: EMBL Grenoble Outstation, France

Structural and functional studies of human ADAM 12 in myoblast fusion and Ebola virus VP40 in assembly

Joanna TIMMINS

A thesis submitted in partial fulfilment of the requirements of the Open University for the degree of Doctor of Philosophy

April 2002

Sponsoring Establishment: NIMR, Mill Hill, U.K.

Collaborating Establishment: EMBL Grenoble Outstation, France

Abstract

Part I:

ADAM 12 is a member of the growing ADAM protein family that displays a conserved domain organisation: a signal sequence, a pro-, a metalloprotease, a disintegrin, a cysteine-rich, a transmembrane and a cytoplasmic domain. ADAM proteins are found in higher eukaryotes and in most mammalian tissues, and appear to be involved in very diverse developmental processes. In this work, a number of biochemical and biophysical techniques were used to obtain structural information on the disintegrin and cysteine-rich domains of ADAM 12 for which no structural data is available. We also investigated the potential role of ADAM 12 in myogenesis and found that ADAM 12 could specifically be retained, by an unknown mechanism, in a perinuclear compartment when expressed in muscle cells, and was unable to provoke myogenesis by itself. A dimerisation partner of ADAM 12 was isolated, namely ADAM 19, which could specifically bind to ADAM 12 *in vitro*.

Part II:

The Ebola virus matrix protein, VP40, is targeted to the plasma membrane where it is thought to induce assembly and budding of virions. Ebola virus VP40 is a monomer in solution consisting of two domains. Cellular localisation studies of different VP40 mutants showed that the C-terminal domain is required for plasma

membrane association. Moreover, we showed that wild type VP40 and an N-terminally truncated form VP40 (31-326) are both released into the cell culture supernatant, when expressed in 293T cells, in vesicular structures that resemble virus-like particles as detected by electron microscopy. These results suggest that VP40 is sufficient for virus assembly and budding. Moreover, VP40 contains two motifs at its N-terminus, which are required for binding to an ubiquitin ligase (Nedd4) and to an inactive E2 enzyme (Tsg101), both of which are here shown to interact with VP40 *in vitro*. These interactions are believed to improve the budding efficiency.

Acknowledgements

I would like to thank EMBL and Dr. S. Cusack for giving me the possibility to do my PhD work at EMBL, in a very ambitious and stimulating scientific environment.

I am very grateful to my director of studies, Dr. W. Weissenhorn, for all his help, technical expertise and guidance throughout my PhD, but also for the time spent reading the thesis drafts.

Many thanks to my second supervisor, Dr. R. Ruigrok, for his useful discussions and advice during the course of my PhD and for proofreading my thesis.

I would also like to thank my University supervisor, Prof. Sir J. Skehel, for devoting his precious time to this work and for his encouragements and fruitful discussions during my visits to the NIMR.

Thanks a lot to Kristine Cronhelm and Dr. R. King for their help with all the paperwork and the organisation of the examination panel for the Open University.

I am also very grateful to Sylvie Ricard-Blum for all her help and patience with the BIAcore experiments.

I would also like to say thank you to all the people at the EMBL Outstation and more particularly to the members of my lab: Andreas, Vassiliy, Maria, Jan and Stéphanie, and a special thanks to Sandra for her technical help but more importantly for her kindness and support.

Many thanks to Guy for all the electron microscopy work.

Thanks to Martine, Serge, Carole, Delphine, Cathy, Fabrice and Annie for providing a pleasant (and sometimes very amusing) work environment.

Last, but not least, I am extremely grateful to my family and friends for their continuous moral support and encouragements.

List of Contents

List of Tables & Figures

Tables:

Figures:

Abbreviations

ADAM: A disintegrin and metalloprotease

BLAST: Basic local alignment search tool

bp: Base pair

kbp: Kilo base pair

CD: Circular dichroism

CFP/GFP: Cyan/Green fluorescent protein

kDa: Kilo dalton

DMEM: Dulbecco's modified Eagle's medium

DNA: Deoxyribonucleic acid

cDNA: Complementary DNA

dNTPs: Deoxyribonucleoside triphosphates

DTT: Dithiothreitol

EDTA: Ethylenediaminetetraacetic acid

EGF: Epidermal growth factor

ER: Endoplasmic reticulum

EST: Expressed sequence tag

FITC: Fluorescein-isothiocyanate

FPLC: Fast performance liquid chromatography

GP: Glycoprotein

GST: Glutathione-S-Transferase

HECT: Homologous to E6-AP C-terminus

His: Histidine

HIV: Human immunodeficiency virus

HTLV-II: Human T-cell leukaemia virus type II

IPTG: Isopropyl-b-D-thiogalactopyranoside

min: Minute

mg: Milligramme

μg: Microgramme

M: Molar

mM: Millimolar

μM: Micromolar

nM: Nanomolar

pmol: Picomoles

ml: Millilitre

μl: Microlitre

nm: Nanometre

MBP: Maltose binding protein

MOI: Multiplicity of infection

MVB: Multivesicular body

MW: Molecular weight (kDa)

NP: Nucleoprotein

NTA: Nitrilotriacetic acid

RNP: Ribonucleoprotein

OD: Optical density

PAGE: Polyacrylamide gel electrophoresis

SDS-PAGE: Sodium dodecyl sulfate-PAGE

PBS: Phosphate buffered saline

PCR: Polymerase chain reaction

RT-PCR: Reverse transcriptase-PCR

PEG: Polyethylene glycol

Pfu: Plaque forming unit

pI: Isoelectric point

PM: Plasma membrane

PMSF: Phenylmethylsulfonyl fluoride

RNA: Ribonucleic acid

mRNA: Messenger RNA

tRNA: Transfer RNA

rpm: Revolutions per minute

RSV: Rous sarcoma virus

SH3: Src homology 3

SIV: Simian immunodeficiency virus

SVMP: Snake venom metalloprotease

TCA: Trichloroacetic acid

N-terminus: Amino-terminus

C-terminus: Carboxyl-terminus

TEV: Tobacco etch virus

TGN: *trans*-Golgi network

UEV: Ubiquitin enzyme variant

UV: Ultra-violet

VEEV: Venezuelan equine encephalitis virus

VSV: Vesicular stomatitis virus

X-gal: 5-bromo-4-chloro-3-indolyl-b-D-galactoside

Summary of Parts I and II

Membrane reorganisations such as membrane fusion and virus budding events require protein machines to ensure specificity and to generate the lipid perturbations which will lead to fusion or fission of otherwise stable membrane bilayers. The identification and characterisation of proteins involved is therefore essential to understand these processes.

ADAM 12 is a member of the growing ADAM (A Disintegrin And Metalloprotease) protein family that displays a conserved domain organisation: a signal sequence, a prodomain, a metalloprotease, a disintegrin, a cysteine-rich, a transmembrane and a cytoplasmic domain. ADAM proteins are found in higher eukaryotes and in most mammalian tissues, and are involved in very diverse developmental processes. At the time of its discovery, ADAM-12 had been directly implicated in cell-cell fusion events by potentially acting as fusion protein, similar to viral fusion proteins. The goal of my work was the structural analysis of the extracellular domains of ADAM 12 containing a putative fusion peptide. In this work, a number of biochemical and biophysical techniques were used to obtain structural information on the disintegrin and cysteine-rich domains of ADAM 12 for which no structural data was available. Finally, protein crystals of the disintegrin and cysteine-rich domains were obtained, but these failed to diffract X-rays to high resolution and trials to improve the quality of the crystals did not succeed. In view of the difficulties encountered with the structural work I also pursued a functional approach to elucidate the role of ADAM-12 in myoblast fusion. I found that ADAM 12 could specifically be retained, by an unknown mechanism, in a perinuclear compartment when expressed in muscle cells, but was unable to provoke myogenesis by itself. A dimerisation partner of ADAM 12

was also isolated, namely ADAM 19, which could specifically bind to ADAM 12 *in vitro* and may thus improve future crystallization trials.

In parallel to this functional work on ADAM 12, I also studied the Ebola virus matrix protein VP40, which is a key component of virus assembly and budding. The laboratory techniques acquired during my work on the ADAM proteins immediately led to the finding that the C-terminal domain of VP40 is required for plasma membrane association *in vivo* and VP40 is released into the cell culture supernatant in vesicular structures that resemble virus-like particles as detected by electron microscopy. These results suggested that VP40 is sufficient for virus assembly and budding. Moreover, VP40 contains two motifs at its N-terminus, which are required for binding to two cellular factors, an ubiquitin ligase (Nedd4) and to an inactive E2 enzyme (Tsg101), both of which are shown to interact with VP40 *in vitro*. These interactions are believed to be crucial for virus budding.

Overall, the structural and functional characterisation of ADAM-12 indicates that it is not directly involved in membrane interaction processes such as fusion of myoblast cells. In addition, I have shown that the matrix protein VP40 interacts with membranes and is sufficient for producing membrane enveloped virus-like particles. Furthermore the structural and functional analysis of VP40 complexed with cellular factors will help to elucidate the role of these protein machines in virus assembly and budding processes, which ultimately involve a membrane fission step to guarantee the integrity of the cellular and the newly formed viral membranes.

Part I

Structural and Functional Studies of Human ADAM 12 in Myoblast Fusion

Chapter 1: Introduction

1. Cell-cell fusion

1.1. Cellular membrane fusion events

Membrane fusion is a ubiquitous cell biological process. These fusion events can be separated into three main classes: fusion resulting from housekeeping functions such as intracellular vesicle transport and exocytosis, virus-cell fusion and intercellular fusion events that occur between specialised cells, such as gametes and myoblast cells. Fusion reactions can be categorised into two types according to their topology (White, 1992). In many fusion events, the leaflets that face the cytoplasm make the initial contact; this category encompasses fusion of intracellular macromolecular carrier vesicles with their target organelles. This type of fusion is known as endoplasmic fusion, in contrast to exoplasmic fusion which involves the exoplasmic leaflets of membranes, as is seen for cell-cell and virus-cell fusion reactions. Given the differences between the extracellular/luminal and the cytoplasmic environments, the proteins and mechanisms involved in endo- and exoplasmic fusion events may be very different. The two membrane surfaces have different lipid and protein compositions (White and Blobel, 1989).

Compared with our knowledge of virus-cell fusion, information about cell-cell fusion remains at a rudimentary level. So far, two general principles have guided the work on cell-cell fusion. The first is that, like all biological membrane

fusion reactions, cell-cell fusion events are energetically unfavourable and will, therefore, require protein-protein interactions in order to be specific and overcome the energetic barrier. The second is that as cell-cell fusion events are topologically equivalent to virus-cell fusion events, the proteins that mediate cell-cell fusion reactions may share traits with viral fusion proteins.

Cell-cell fusion events are critical to the development of multicellular organisms; fusion of gametes to form the zygote, myoblasts to form myotubes and monocytes to form osteoclasts, all set off new developmental pathways. Current efforts are, therefore, focused on identifying cell-cell fusion proteins. This is a much more difficult task than it was for the virus-cell fusion proteins for two major reasons. First, cells have many more surface glycoproteins, and second, cell-cell fusion events are preceded by cell commitment and cell-cell adhesion (and in some cases cell migration), making it challenging to determine whether a protein implicated in fusion is involved in cell determination, migration, adhesion, or in the actual fusion mechanism itself (Hernandez et al., 1996). Two major approaches have been undertaken to identify candidate cell-cell fusion proteins: (a) characterisation of antigens recognised by antibodies that block fusion without noticeably blocking cell commitment or adhesion (Primakoff et al., 1987; Saginario et al., 1995), and (b) characterisation of mutants that are defective in cell-cell fusion (Ferris et al., 1996; Trueheart and Fink, 1989).

1.2. Myoblast fusion

Of the known cell-cell fusion reactions, the two that have been investigated most intensely are myoblast fusion and gamete fusion. The fusion of myoblasts leading to the formation of skeletal muscle is of particular interest. This fusion

event must be very tightly controlled during development if the final muscles are to be patterned and sized correctly (Blau et al., 1993). The ability to influence this process would be of great therapeutic value. Skeletal muscle development (Figure 1) involves the formation of multinucleated myotubes, a process that can be divided into a series of steps: cell determination, cell-cell recognition, adhesion, alignment, and finally membrane fusion (Knudsen and Horwitz, 1978). In the determination step, the myoblasts start to produce the proteins that make the cells competent to fuse. The myoblasts then migrate to an appropriate fusion target, where the cells adhere to each other in a calcium-dependent manner (Knudsen and Horwitz, 1977). After adhesion, the cells align along their long axis, bringing their plasma membranes into close apposition, leading to local membrane fusion events.

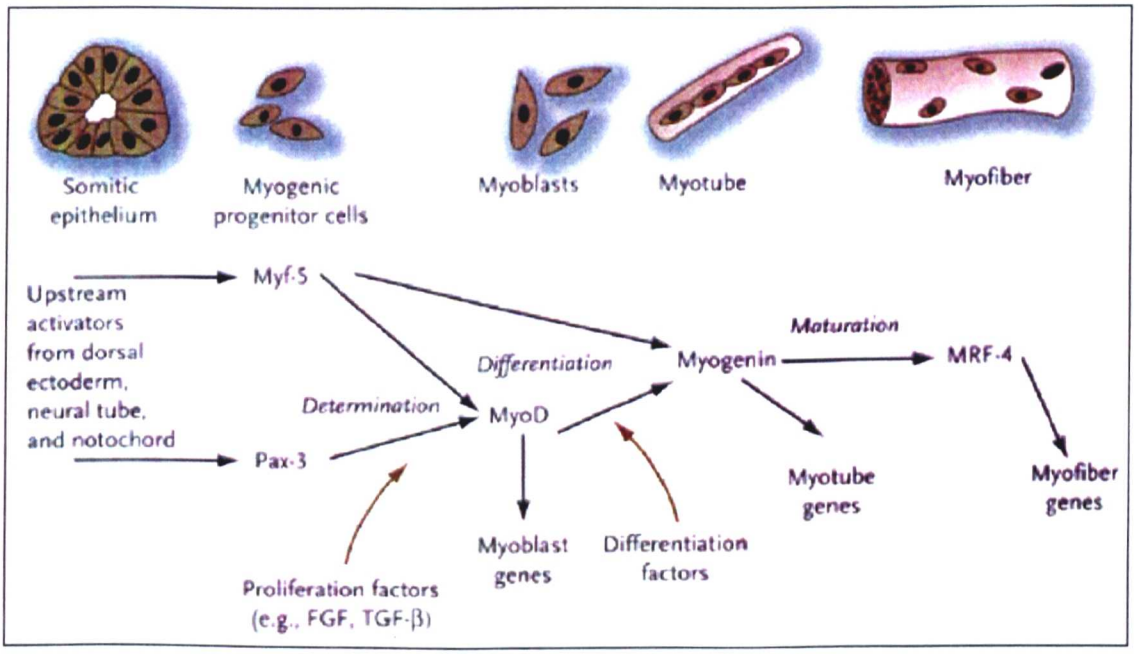

Figure 1: Schematic diagram of the multiple steps involved in myoblast development and the regulatory genes that coordinate skeletal muscle differentiation. Taken from:

http://www.tulane.edu/~embryo/Lectures%20Fall%202000/08Muscle/Muscle%20Development.pdf

The molecules involved in myoblast differentiation are very diverse (Table 1): growth factors, transcription factors, ion channels, cell-surface receptors, components of the extracellular matrix etc. Most of the original work on myoblast fusion was done on *in vitro* systems in which cultures of myoblasts can be synchronised for fusion by varying the calcium concentration, and this work allowed the identification of new proteins involved in fusion (Knudsen, 1992). Notably the use of the mouse C2C12 cell line and the human rhabdomyosarcoma (RD) cell line have led to the discovery of many molecules that are essential for myoblast fusion. This is the case for various membrane proteins such as N- and M-cadherins (Goichberg and Geiger, 1998) or integrins (Rosen et al., 1992), which have been shown to participate in myotube formation. Nonetheless, the environment and morphology of cultured myoblasts differ greatly from the *in vivo* state, and many crucial questions remain unanswered. What steps are required at the molecular level to enable pairs of myoblasts to align and fuse? Which proteins and other molecules mediate these successive steps, and how do they interact with each other?

Table 1: Summary of the various molecules known to be involved in myoblast development.

Name	Localisation	Function	Reference
Growth factors: Wnt, Sonic hedgehog, basic Fibroblast Growth Factor (bFGF), Transforming Growth Factor (TGF-β1), Insulin-like Growth Factor (IGF I & II)	Extracellular	Several growth factors interact synergistically to promote somite myogenesis	(Pirskanen et al., 2000)
Transcription factors: MyoD family (MyoD, Myf5, myogenin, MRF4) MEF2 family	Nuclear/ cytoplasmic	Transcriptional control of myogenesis	(Megeney and Rudnicki, 1995) (Black and Olson, 1998)
Lame duck	Nuclear/ cytoplasmic	Transcriptional regulator of MEF2 in fusion-competent myoblasts	(Duan et al., 2001)

Histone acetyltransferases (HATs) & deacetylases (HDACs)	Nuclear	Control the activation and repression of muscle-specific genes by associating with MEF2 proteins	(McKinsey et al., 2001)
Nesprins	Nuclear membrane; changes during differentiation	Proposed to maintain nuclear organisation and structural integrity	(Zhang et al., 2001)
Myoblast city	Cytoskeletal protein	Involved in Rho/Rac GTPase signalling pathway	(Erickson et al., 1997)
Blown fuse	Cytoplasmic	Required for progression beyond the prefusion complex stage	(Doberstein et al., 1997)
Antisocial	Cytoplasmic	Acts as an adaptor protein; interacts with the cytoplasmic domain of dumbfounded and with myoblast city	(Chen and Olson, 2001)
Muscle-specific RING-finger protein (MURF)	Cytoplasmic	Acts as a regulator of the microtubule (MT) network of striated muscle cells and reveals a link between MTs organisation and myogenesis	(Spencer et al., 2000)
Muscle-specific β1-integrin-binding protein (M1BP)	Cytoplasmic	Binds to the proximal cytoplasmic region of β1-integrins; controls progression of muscle differentiation	(Li et al., 1999)
Kir 2.1 (K+ ion channel)	Cell surface	Functional Kir2.1 channels, responsible for hyperpolarisation of myoblasts, are required for human myoblast fusion	(Fischer-Lougheed et al., 2001)
T-type Ca2+ channels	Cell surface	Responsible for the increase in intracellular Ca2+ concentration observed at the onset of myoblast fusion	(Bijlenga et al., 2000)
β1-Integrins	Cell surface	Regulates the expression of muscle-specific genes	(Rohwedel et al., 1998)
Syndecans	Cell surface	Role in regulation of skeletal muscle terminal differentiation	(Fuentealba et al., 1999)
N-Cadherin	Cell surface; cell-cell adherens-type junctions	Role in cell-cell adhesion and in signalling events to activate myogenic differentiation	(Goichberg and Geiger, 1998)
β-Catenin	Cell-cell junctions	Activates early steps of myogenic differentiation	(Goichberg et al., 2001)
Vascular cell adhesion molecule (VCAM-1)	Cell surface	Role in secondary myogenesis; interacts with integrin alpha 4 (VLA-4), leading to alignment of secondary myoblasts	(Rosen et al., 1992)
Integrin VLA-4	Cell surface		
Laminin/ fibronectin/ vitronectin	Extracellular matrix	Promote adhesion, migration and proliferation of mammalian myoblasts; laminin interacts in myoblasts with integrin α7β1	(Crawley et al., 1997) (Gullberg et al., 1995)

ADAM 12 (meltrin-α)	Cell surface	Putative roles in adhesion, degradation and fusion in myoblasts	(Yagami-Hiromasa et al., 1995)
Transmembrane 4 superfamily (TM4SF) proteins: CD9 and CD81	Cell surface	CD9 can interact with b1-integrins; TM4SF proteins appear to promote muscle cell fusion and support myotube maintenance	(Tachibana and Hemler, 1999)
BOC/CDO receptor complex	Cell surface	Mediates some of the cell-cell interactions between muscle precursors that are required for myogenesis	(Kang et al., 2002)
Dumbfounded	Cell surface (?) of founder cells	Acts as a myoblast attractant, leading to aggregation and fusion of myoblasts	(Ruiz-Gomez et al., 2000)
Sticks and stones (SNS)	Cell surface of fusion-competent myoblasts	Unknown; may possibly be an interacting partner of dumbfounded	(Bour et al., 2000)
Rolling stone	Cell surface	Involved in the fusion process during myogenesis	(Paululat et al., 1997)
Rolling pebbles	Not known	Required in muscle precursor cells and is essential to recruit fusion-competent myoblasts for myotube formation	(Rau et al., 2001)
Hibris	Not known	Proposed to be an extracellular partner of dumbfounded and potentially mediates the response of myoblasts to this attractant	(Dworak et al., 2001)
D-Titin	Cell-surface	Functions to maintain myotube structure and morphology; Rolling pebbles recruits D-Titin in response to the attractant, dumbfounded	(Menon and Chia, 2001)
Gelatinase-B (Matrix Metalloproteinase-9)	Extracellular matrix	Secretion of MMP-9 is involved in the regulation of the matrix turnover, affecting events leading to myotube formation, such as migration	(Lewis et al., 2000)
Neuregulin	Not known	Role as a potent differentiation-promoting activity in membrane fusion and expression of myosin heavy chain	(Kim et al., 1999)
LIM proteins	Muscle attachment sites; Z-bands of striated muscle	Structural role as a component of muscle cyto-architecture	(Stronach et al., 1999)

Drosophila melanogaster is an excellent organism for the study of muscle development. Like skeletal muscles in vertebrates, *Drosophila* muscles consist of syncitial fibres formed by fusion of myoblasts. But in contrast to vertebrates, in *Drosophila* larvae, each muscle is a single myotube. Because muscle development in *Drosophila* is rapid and well documented, it has been possible to identify several genes essential for fusion. In addition, classical genetic mutant analysis has been a very powerful and specific tool for the identification of proteins involved in developmental and cell biological processes. To date several mutants have been described with specific defects in myoblast fusion (Paululat et al., 1999). According to their phenotype, these mutant genes were named *myoblast city* (*mbc*) (Rushton et al., 1995), *blown fuse* (*blow*) (Doberstein et al., 1997) and *rolling stone* (*rost*) (Paululat et al., 1995). Concerning Mbc, it has been proposed that it is not directly involved in fusion, but rather in the cytoskeletal rearrangement process prior to cell adhesion, which renders myoblasts competent for fusion (Erickson et al., 1997). The cytoplasmic Blow protein has been shown to mediate the step between prefusion complex formation and plaque formation. Finally, the integral membrane protein, Rost, appears to act downstream of Mbc and is indispensable for myotube formation. More recently, a new gene was identified, *dumbfounded* (*duf*), which has been shown to encode a member of the immunoglobulin superfamily of proteins that is an attractant for fusion-competent myoblasts (Ruiz-Gomez et al., 2000). Progressively new molecules involved in very diverse pathways leading to myoblast formation are being characterised, but the molecular mechanisms of muscle cell fusion still remain poorly understood.

1.3. Gamete fusion

Our understanding of the mechanisms involved in mammalian fertilisation, however, is a little more advanced, and it can reasonably be assumed that some principles that guide gamete fusion also apply to muscle cell fusion. Interestingly, antibody inhibition studies have implicated several sperm surface proteins in binding and fusion, one of which was named fertilin (Primakoff et al., 1987). Characterisation of this molecule revealed that it possessed a disintegrin domain (a putative integrin binding domain) and was soon recognised to be a member of a large protein family, which displayed a high homology to snake venom metalloproteases. A search for homologous proteins expressed specifically in myoblasts, led to the isolation of three gene products, namely: meltrin-α, β and γ (Yagami-Hiromasa et al., 1995). Interestingly, meltrin-α and meltrin-β were specifically expressed in muscle at embryonic and neonatal stages, while meltrin-γ was expressed ubiquitously. Moreover, meltrin-α was also expressed in a mouse myoblast cell line, C2, and the level of meltrin-α messenger RNA (mRNA) increased dramatically in response to the induction of differentiation, suggesting that meltrin-α is associated with early stages of myotube formation. The finding that a given protein family, namely the ADAM (A Disintegrin And Metalloprotease) family, is involved in both gamete fusion (fertilin) and in myoblast fusion (meltrin-α), suggests that these proteins may be crucial for the developmental process leading to cell-cell fusion.

2. The ADAM protein family

2.1 Characteristics of the ADAM protein family

Together with snake venom metalloproteases (SVMPs) the ADAM proteins make up the reprolysin family of zinc metalloproteases. The ADAM proteins share 30% sequence identity with the soluble SVMPs, which were the first members of this family to be identified (Blobel et al., 1992; Wolfsberg et al., 1995). SVMPs have been extensively studied; full-length SVMPs are processed to generate a metalloprotease, which is able to degrade proteins of the basement membrane such as type IV collagen, and laminin (Hite et al., 1994), and a disintegrin domain, which can inhibit the function of platelets by interacting with platelet integrin αIIbβ3 (Paine et al., 1992). Like the SVMPs, the ADAMs are multidomain proteins, consisting of a signal sequence, a prodomain, a metalloprotease domain, a disintegrin domain and a cysteine-rich domain (Figure 2). In addition, many ADAMs possess several additional C-terminal domains compared to the structure of the SVMPs. Notably, they share an epidermal growth factor repeat (EGF domain) and a transmembrane domain, followed by a cytoplasmic tail. Moreover, some ADAMs have a putative fusion domain in their cysteine-rich domains (Blobel et al., 1992), or instead of the C-terminal domains some ADAMs have a thrombospondin motif (Kuno et al., 1997). ADAMs were first identified in guinea pig sperm, but have since been found in a wide range of higher eukaryotes, such as *C. elegans, Drosophila, Xenopus,* and in most mammalian tissues. However, none have been reported so far in unicellular eukaryotes, bacteria or plants.

ADAMs, like matrix metalloproteases, are produced as zymogens, resulting from the formation of an intramolecular complex between a single cysteine

residue in its prodomain and the essential zinc ion in the catalytic domain, a complex that blocks the active site.

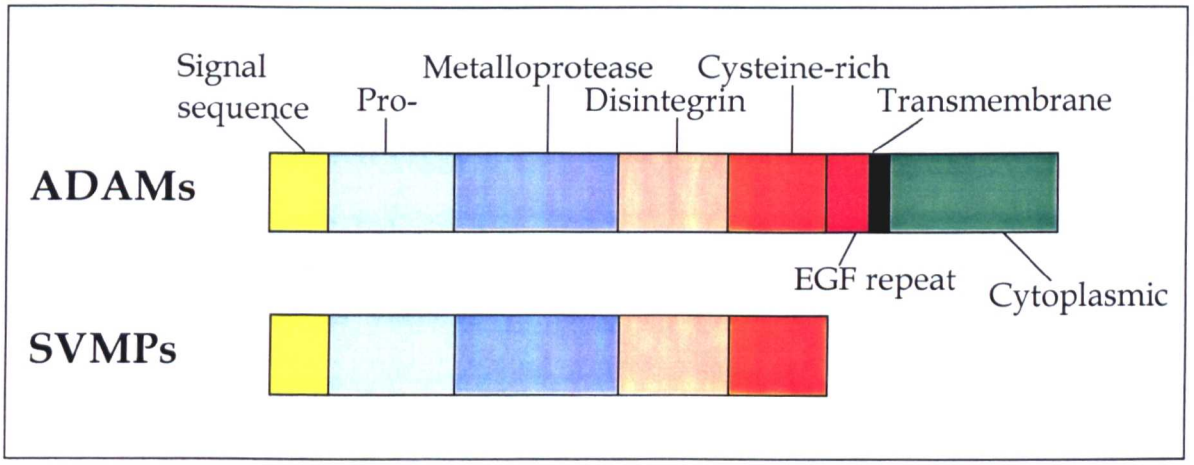

Figure 2: Comparison of the domain organisation of ADAMs and SVMPs.

The latent proteins can be activated by dissociation of the cysteine or enzymatic cleavage of the prodomain. This activation mechanism is called the "cysteine switch" (Van Wart and Birkedal-Hansen, 1990). In some ADAMs the prodomain cleavage sites are readily identified as targets for the ubiquitous furin proteases. Although all ADAMs possess a metalloprotease domain, the zinc-dependent protease consensus sequence **HE**XG**H**XX**G**XX**H**D is not present in all ADAM primary sequences, suggesting that only a subset of the ADAMs actually possesses an active metalloprotease domain (Black and White, 1998). So far based on the amino acid sequence, the following ADAMs (across all known orthologs) display the consensus active site sequence: ADAMs 1, 8-10, 12, 13, 15, 16, 17, 19, 20, 21, 24, 25, 26, 28, 30, and 33 (Primakoff and Myles, 2000).

The disintegrin domains of the ADAMs are believed to be ligands for integrins or other receptors. Normally, disintegrins (as in SVMPs) interact with their substrates through a disintegrin loop that contains the integrin-binding

sequence Arg-Gly-Asp (RGD) at its tip (Kratzschmar et al., 1996). However, with the exception of ADAM 15, also known as metargidin, none of the ADAMs expresses the RGD sequence. Approximately half of the ADAMs contain a negatively charged residue at the position of the aspartic acid of the RGD tripeptide, and it has been suggested that this negatively charged residue is critical for integrin binding (Wolfsberg et al., 1995). In addition, within the ADAM family, the disintegrin loop is variable in length. Different ADAMs may, therefore, interact with different integrins and/or other cell surface receptors. So far, eight different ADAM-integrin interactions have been described: ADAM 15 and ADAM 23 with $\alpha_v\beta_3$ (Cal et al., 2000; Zhang et al., 1998), ADAM 15 with $\alpha_5\beta_1$ (Nath et al., 1999), ADAM 15 and ADAM 12 with $\alpha_9\beta_1$ (Eto et al., 2000), ADAM 9 with $\alpha_v\beta_5$ (Zhou et al., 2001), and ADAM 2 and ADAM 9 with $\alpha_6\beta_1$ (Chen and Sampson, 1999; Nath et al., 2000).

The cysteine-rich domain is characterised by the presence of a high percentage of cysteinyl residues displayed in a distinctive sequence pattern. The function of this domain remains unclear; it has been suggested to function as a protein-protein interaction domain (Iba et al., 2000). Moreover, in a few ADAMs (ADAMs 1, 5 9, 11, 12 20 and 21) the cysteine-rich domain contains a putative fusion peptide, which like viral fusion peptides may promote membrane fusion (Blobel et al., 1992). These candidate fusion peptides were identified by their overall hydrophobicity and their ability to model as a α-helix with a very hydrophobic side. Nonetheless, extensive mutagenesis and biochemical analyses will be needed to define the real role of these potential fusion peptides (Hooft van Huijsduijnen, 1998).

While the transmembrane domain functions as a membrane anchor for the ADAMs, the presence of an EGF-repeat and a cytoplasmic tail suggests that these proteins may serve as signal transducers between the extracellular and intracellular space. The cytoplasmic tails are very variable in length and do not share sequence similarity with one another or with other proteins. Some ADAMs possess proline-rich sequences in their cytoplasmic tails, which exhibit a conserved potential Src homology 3 (SH3) ligand domain RPXPXXP.

The presence of these multiple domains in a single polypeptide chain suggests that the ADAM proteins are capable of four potential functions: proteolysis, adhesion, signalling and maybe fusion. Currently, there is evidence that each of these domains (with the exception of the transmembrane domain) has a functional, and not just a structural, role in at least one ADAM.

2.2 Multiple roles of the ADAM family members

The number of genes known to be members of the ADAM family has grown rapidly and the biological function of most of these new members seems to be unclear. But in view of their multiple domains, they could potentially play various and complex roles for the development and maintenance of an organism.

The first ADAMs, ADAMs 1 and 2 (also referred to as fertilin α and β) were isolated during an analysis of a sperm protein complex involved in mammalian fertilisation (Primakoff et al., 1987). ADAM 1 and 2 form a heterodimeric complex (fertilin), which assembles during biosynthesis in testis, and is processed in at least two steps (Blobel et al., 1990). There is now good evidence for the involvement of this complex in fertilisation. Monoclonal antibodies against fertilin can inhibit sperm-egg fusion (Primakoff et al., 1987). Peptides corresponding to the

disintegrin domain of fertilin also inhibit sperm-egg fusion in guinea pig (Myles et al., 1994) and in mouse (Almeida et al., 1995). In addition, an $\alpha_6\beta_1$ integrin on mouse eggs acts as a receptor for the sperm (Almeida et al., 1995). Moreover, mice lacking ADAM 2 were shown to be deficient in sperm-egg membrane adhesion, sperm-egg fusion, migration from the uterus to the oviduct, and binding to the egg zona pellucida (Cho et al., 1998). Nevertheless, it should be emphasised that mammalian fertilisation involves much more than the interaction of ADAM 2 on sperm and $\alpha_6\beta_1$ on eggs. In fact many results suggest the involvement of several sperm ligands and egg receptors in gamete membrane interactions (reviewed in (Evans, 2001)). Additional ADAMs have been found in the testis, many of which are specifically expressed in testis; this is the case for the following ADAMs: 2, 3, (5), (6), 16, 18, 20, 21, 24, 25, 26, 29, and 30. There is now evidence that ADAM 3 (also known as cyritestin) in mouse may play a significant role in fertilisation since synthetic peptides corresponding to this protein strongly inhibit both sperm-egg binding (Yuan et al., 1997) and fertilisation. But for most of these testis-specific ADAMs, a role in reproductive functions has not yet been addressed. In addition, there are numerous other types of gamete molecules (such as complement components or sulfated glycolipids), less well characterised on gametes but with potential to be involved in these processes (Evans, 1999). The detailed mechanism of mammalian fertilisation, at a molecular level, thus still remains largely unknown, but exemplifies the complexity of cell-cell fusion processes, in which the ADAM proteins are clearly involved.

Besides their role in cell differentiation and fusion, some ADAMs are catalytically active via their metalloprotease domain in critical physiological functions. The first observation was that human ADAM 17 (also known as TACE,

tumour necrosis factor-α-converting enzyme) is responsible for the release of the membrane-anchored cytokine tumour necrosis factor-α (TNF-α) from the plasma membrane (Black et al., 1997). Bovine and human ADAM 10 have also now been shown to possess a TNF-α converting enzyme activity (Lunn et al., 1997). TNF-α is a cytokine that is normally produced in response to infection or injury, but can also contribute to a variety of inflammatory disease states, such as rheumatoid arthritis and Crohn's disease. Since ADAM 10 and ADAM 17 are able to cleave the active ectodomain of TNF-α from its membrane-anchored precursor, these molecules are believed to play a critical role in the control of this shedding process. As a result, ADAM 10 and ADAM 17 are often referred to as "sheddases" (Primakoff and Myles, 2000). In addition, the second observation was that ADAM 10 in *Drosophila* (known as Kuzbanian) is able to cleave the extracellular domain of NOTCH, a cell surface receptor that is involved in signal tansduction (Pan and Rubin, 1997). It seems that ADAM 10 plays an evolutionarily conserved role in NOTCH processing (Black and White, 1998). In *C. elegans* ADAM 10/SUP-17 (homologue of Kuzbanian) and LIN-12 (homologue of NOTCH) both determine cell fates during development of the hermaphrodite gonad and development of the vulva (Wen et al., 1997). ADAM 10 in *Xenopus* is also believed to play the same role in neurogenesis. Finally, a third ADAM, ADAM 9, has also been shown to act as a sheddase. ADAM 9 can shed the heparin-binding EGF-like growth factor. Both the membrane-anchored and soluble form of this growth factor are active, but the soluble, diffusible form can act on cells distant from the site of its shedding (Lammich et al., 1999). The other 14 predicted active ADAM metalloproteases are still lacking an identified endogenous substrate.

On the basis of the known functions of only a few ADAM proteins, potential roles for other ADAM family members have been proposed. Elucidation of mechanisms that regulate the adhesion properties of the ADAMs, and understanding of whether disintegrin-mediated adhesion can lead to the activation of intracellular signalling pathways via the cytoplasmic tail, should prove interesting. Moreover, a better understanding of what determines the substrate specificity of these metalloproteases and how the catalytic activity is regulated are also of great importance. Besides ADAM 1 and 2, which have been the focus of many studies over the past decade in an attempt to shed light on mammalian fertilisation, another member of the ADAM family, namely ADAM 12, has attracted much attention.

3. ADAM 12 protein

3.1 Initial characterisation of ADAM 12

The proteins of the ADAM family have a widespread cell distribution including sperm, epididymis, epithelium, placenta, ovary, breast, skeletal muscle, heart, liver, kidney, small intestine, colon, brain, thymus, spleen, lung, bone, monocytes, macrophages, and leukocytes (Wolfsberg and White, 1996). In 1995, Yagami-Hiromasa *et al.* reported the identification of three new ADAM family members (meltrin-α, β and γ). These new genes were isolated in a search for fertilin-related genes expressed in muscle. Complementary DNAs (cDNAs) prepared from a mouse myogenic cell line were amplified using degenerative primers for conserved sequences in ADAM 1 and 2. Meltrin-α and β were specifically expressed in muscle and bone, while meltrin-γ was ubiquitously

expressed. Interestingly, the levels of meltrin-α mRNA increased dramatically in response to the induction of differentiation, suggesting that meltrin-α is associated with the initial stages of myotube formation (Yagami-Hiromasa et al., 1995). Meltrin-α has since been further characterised and is a member of the ADAM family, and according to the new nomenclature was named ADAM 12. Initial studies on the possible role of mouse ADAM 12 in C2 cells revealed that the expression of the anti-sense RNA suppressed ADAM 12 induction, leading to a block in myotube formation. Overexpression studies, using both the full-length ADAM 12 (in its latent state) and a truncated form (activated state), missing the prodomain and most of the metalloprotease domain, showed that expression of the truncated form facilitated fusion of C2 cells, while overexpression of the wild-type protein suppressed fusion (Yagami-Hiromasa et al., 1995). Nonetheless, when non-muscle cells were made to express the active, truncated form of ADAM 12, they did not acquire the ability to fuse, adhere to each other, or undergo myogenesis, suggesting that ADAM 12 alone cannot induce these events.

More recently, it was shown that mouse ADAM 12 was also expressed in primary osteoblast-like cells, in addition to myoblasts (Inoue et al., 1998). Moreover, it was shown that ADAM 12 is involved in the development of giant cells and osteoclasts from their respective precursor cells (Abe et al., 1999). Interestingly, expression of ADAM 12 was induced under conditions that promote giant cell formation and interference of the expression of this construct by antisense oligonucleotides resulted in inhibition of multinucleation. However, further studies to look at the expression pattern of ADAM 12, revealed that by using a much more sensitive method for studying gene expression (i.e. reverse-transcriptase-polymerase chain reaction (RT-PCR), as opposed to Northern

blotting in the case of Yagami-Hiromasa's studies), ADAM 12 expression could be detected in a wide variety of tissues (skeletal muscle, lung, testis, uterus, kidney, and intestine), many of which consist of non-fusagenic cells (Harris et al., 1997). These results also suggest that ADAM 12 may not function exclusively in processes involving fusion.

The human homologue of meltrin-α, hADAM 12, was cloned in 1998 (Gilpin et al., 1998). Interestingly, some ADAM mRNAs appear to undergo alternative splicing events producing shorter, secreted forms. This was found to be the case for hADAM 12. Cloning of hADAM 12 from a human placenta cDNA library revealed that it had two alternatively spliced versions of the same gene: a full-length protein (ADAM 12-L) and a truncated form (ADAM 12-S) lacking the transmembrane and cytoplasmic domains (Figure 3).

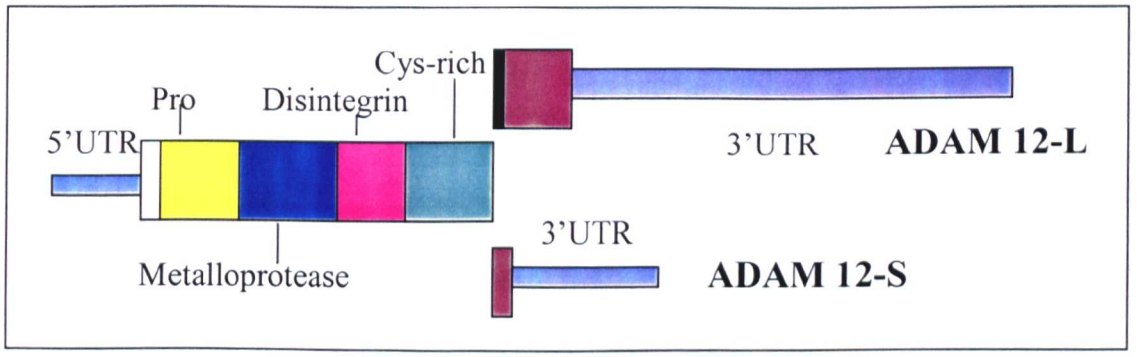

Figure 3: Schematic diagram of the hADAM 12-L and hADAM 12-S mRNAs (adapted from (Gilpin et al., 1998)).

ADAM 12-S has only been detected in placenta, whereas ADAM 12-L can be found in cardiac, smooth and skeletal muscle in addition to placenta. Both forms, however, were found to be expressed in several cultured human cell lines, such as the RD (rhabdomyosarcoma) cell line. Gilpin *et al.* found that cells transfected with ADAM 12-S appear to be very potent in provoking myogenesis *in vivo*. This was

shown by transfecting a human embryonal rhabdomyosarcoma cell line, A204 (that does not express ADAM 12), with a construct corresponding to the disintegrin, cysteine-rich domains and the unique carboxyl terminus of the ADAM 12-S. The A204 cell line is unable to differentiate *in vitro*, whether it be spontaneously or after transfection with the ADAM 12-S construct. However, nude mice tumours generated from these ADAM 12-S transfected cells contained a striking pattern of ectopic muscle cell formation. Nonetheless, the mechanism by which ADAM 12-S may be involved in the recruitment and differentiation of muscle progenitor cells is not clear.

3.2 Functional roles of ADAM 12

Sequence analysis of the newly cloned hADAM 12 revealed that it was 81% identical to its mouse homologue. The human ADAM 12 metalloprotease domain contains the conserved zinc-binding motif **HEXGHXXGXXHD** regulated by a potential cysteine switch in the prodomain, and on the basis of this sequence is presumed to be catalytically active. A consensus sequence for cleavage by furin-type endopeptidases is also present at the boundary between the pro- and the metalloprotease domains. The disintegrin domain contains a putative integrin-binding loop, although like most ADAMs, ADAM 12 does not have an RGD sequence, but instead the amino acids Ser-Gln-Ser (SNS) at this position followed by an additional cysteine residue. The cysteine-rich domain of hADAM 12 was suggested to contain a putative fusion peptide and an EGF-like repeat. Finally, the cytoplasmic domain of hADAM 12-L is proline-rich (32 out of 179 amino acids) and contains at least three SH3 binding motifs (RXXPXXP). Over the past years,

hADAM 12 has been the focus of many functional studies in order to clarify its potential role in cell-cell adhesion and fusion.

As expected from the primary sequence analysis, hADAM 12 has been shown to be an active metalloprotease (Loechel et al., 1998). In this work, they demonstrate that hADAM 12 can react with purified α2-macroglobulin, resulting in the appearance of cross-linked monovalent and multivalent complexes that migrate as slower, high molecular weight bands on a SDS-polyacrylamide gel. This reaction could be inhibited by addition of the metalloprotease inhibitor 1, 10-phenanthroline. Moreover, mutation of the active site glutamate residue, present in the zinc-binding site, eliminated the protease activity of ADAM 12.

At present, much effort is aimed at identifying the physiological substrate(s) of the ADAM 12 metalloprotease domain. It was recently shown that purified, recombinant hADAM 12-S could cleave purified insulin-like growth factor binding protein-3 (IGFBP-3) and 5 (IGFBP-5), a cleavage reaction that could also be inhibited by 1, 10-phenanthroline. IGFBP-3 is the major IGF-binding protein in human serum, and is mostly degraded in maternal serum during the course of pregnancy (Loechel et al., 2000). The fact that ADAM 12-S can interact with and cleave IGFBP-3, in addition to being present in human pregnancy serum, and not in nonpregnancy serum, strongly suggests that IGFBP-3 is a candidate physiological substrate for degradation by ADAM 12-S. Moreover, recent data suggests that ADAM 12 may be involved in cellular signalling in cardiomyocytes (Asakura et al., 2002). In this study it was found that shedding of heparin-binding epidermal growth factor (HB-EGF) as a result of metalloprotease activation, and subsequent transactivation of the epidermal growth factor receptor (EGFR), occurred when cardiomyocytes were stimulated by G-protein coupled receptors

(GPCR). The result of this activation is cardiac hypertrophy, which when prolonged typically results in chronic heart failure or sudden cardiac death. ADAM 12-mediated cleavage of HB-EGF appears to trigger EGFR transactivation. Thus both HB-EGF and ADAM 12 constitute potential targets for the treatment of cardiac hypertrophy.

Further studies also showed that ADAM 12 is synthesised as a latent metalloprotease that can be activated chemically by treatment with *N*-ethylmaleimide (NEM), which presumably alkylates cysteine 179 (assumed to be the residue involved in the cysteine switch mechanism, on the basis of sequence comparisons) and destroys the cysteine switch of the prodomain. Mutation of this crucial cysteine 179 also results in a constitutively active metalloprotease domain (Loechel et al., 1999). Physiologically, it is believed that this activation step is achieved by proteolytic cleavage of the prodomain, resulting in unmasking of the catalytic centre, as the cysteine dissociates from the zinc atom present at the catalytic site of the metalloprotease domain. The endogenous furin-type endopeptidase responsible for this cleavage *in vivo* has not yet been determined. Interestingly, activation of ADAM 12-S has also been achieved by a cysteine-switch-independent mechanism (Loechel and Wewer, 2001). It was found that the presence of 50 μM Zn(II) or 10 μM Cu(II) was needed for degradation of α2-macroglobulin by ADAM 12-S. Moreover, a mutant form of ADAM 12-S, in which cysteine 179 was mutated to serine, behaved identically to wild-type ADAM 12-S; both Cu(II) and Zn(II) activated it. These results suggest that the cysteine switch and copper activation are two separate components of the ADAM 12 latency/activation mechanism. The current model, thus suggests that ADAM 12 is maintained in its latent state in the endoplasmic reticulum, and as it passes

through the *trans*-Golgi network, ADAM 12 is cleaved in between the pro- and the metalloprotease domain. However, the propeptide remains bound to the protease (the prodomain coelutes with the protease from a gel filtration, after cleavage), and a second activation step involving copper is then required to obtain a fully activated metalloprotease. Copper binding may result in oxidation of one of the amino acid residues in the binding site, and thereby changes the conformation of ADAM 12-S.

From primary sequence analysis of ADAM 12, a major question that comes to mind is whether the non-RGD (SNS) disintegrin domain of ADAM 12 interacts with integrins. The use of recombinant disintegrin domains and cells expressing recombinant ADAMs helped to address this question. Using CHO cells expressing different recombinant integrins, Eto *et al.* showed that the human and mouse ADAM 12 disintegrin domains could support adhesion of $\alpha_9\beta_1$-CHO cells (Eto et al., 2000). The human ADAM 15, which has the RGD motif, binds to $\alpha_v\beta_3$ in an RGD-dependent manner, and to $\alpha_9\beta_1$ in an RGD-independent manner, as does ADAM 12. Considering a wide distribution of ADAMs and $\alpha_9\beta_1$, ADAM/$\alpha_9\beta_1$-mediated cell-cell interaction may be involved in many developmental and pathological situations, including myoblast fusion.

Analysis of the tissue distribution pattern of ADAM 12 revealed that ADAM 12 expression was upregulated in cancer, and that in several tumours ADAM 12 immunostaining was enriched along the cell surfaces (Iba et al., 1999). In view of these initial results, Iba *et al.* addressed the question as to whether ADAM 12 was involved in cell adhesion. The use of the recombinant cysteine-rich domain of hADAM 12 showed that indeed, carcinoma cells could attach, although not spread, on plates coated with the cysteine-rich domain. Further studies have

now shown that mesenchymal cells can attach, spread and even form focal adhesions and organise stress fibres in response to ADAM 12 (Iba et al., 2000). In addition, this work also revealed that this process was mediated by two types of cell surface molecules: syndecans and integrins. Syndecans also seem to be involved in the process of attachment of carcinoma cells to the cysteine-rich domain of ADAM 12. Syndecans are thus responsible for the cell attachment and β_1 integrins for the subsequent cell spreading. Syndecans are abundant cell surface heparan sulfate proteoglycans (HSPGs) involved in multiple biological processes (Carey, 1997; Woods and Couchman, 1998). In this study, low concentrations of heparin completely eliminated the interaction between mesenchymal cells and the immobilised ADAM 12. Cell lines such as ARH-77, which express very little cell surface heparan sulfate, do not attach to ADAM 12. It, therefore, appears that both the disintegrin and cysteine-rich domains are involved in interactions with cell surface receptors. Most of this work has been done with single recombinant domains, and as a result it is still unclear whether both of these domains can function at the same time. The binding affinity for syndecans, for example, is relatively low (dissociation constants in the range of $2x10^{-8}$-$2x10^{-5}$M) so the binding of the disintegrin domain to an integrin receptor, in addition to that of the cysteine-rich domain to syndecans, may ensure a higher overall binding affinity and substrate specificity. Moreover it has been estimated that integrin receptors extend from the surface of the plasma membrane for a distance of approximately 20 nm (Hynes, 1987), as opposed to a typical heparin sulfate chain that is predicted to form an extended linear polymer with an estimated length of near 60 nm (Brickman et al., 1995). As a result, this would extend the effective range of cellular interactions (Wong et al., 1997) with integrin/syndecan ligands by a factor of 3.

Recently, heparan sulphate proteoglycans have been implicated in the entry of HIV-1 into macrophages. These HSPGs most likely concentrate viruses at the surface of cells, thus increasing the probability of HIV-1 particles to interact with scarce CD4 receptors (Saphire et al., 2001).

Finally, the cytoplasmic tail of ADAM 12 has been the subject of much interest over the past few years. The function of the cytoplasmic domains of ADAMs is much less clear. The cytoplasmic domain of ADAM 12 contains eight copies of the sequence PXXP, the core consensus sequence required for binding to SH3 domains. The minimal consensus binding motifs for the SH3 domain of Src have been determined as RXXPXXP (class I ligands) or PXXPXR (class II ligands). Both of these motifs are present in the cytoplasmic tail of ADAM 12. It has now been shown that the cytoplasmic tail of mouse ADAM 12 can bind to the SH3 domain of Src protein tyrosine kinase both *in vitro* and *in vivo* (Kang et al., 2000). Interestingly, this interaction with the amino-terminal proline-rich region of ADAM 12 results in stimulation of Src tyrosine kinase activity. These initial results provide insight into the possible function of the cytoplasmic domain of ADAM 12, and indicate that ADAM 12 may well play a role in transmembrane signalling. In addition, Suzuki *et al.* investigated the binding of ADAM 12 to a variety of SH3 domain-containing proteins such as: Yes, Abl, Grb2 and the p85 subunit of phosphatidylinositol 3-kinase (PI3K) (Suzuki et al., 2000). In this study they show that, *in vitro*, the cytoplasmic domain of ADAM 12 could bind to recombinant SH3 domains of Yes and Grb2, in addition to that of Src. Grb2 and Src also coimmunoprecipitated with ADAM 12 from both myoblast and myotube cell lysates, suggesting that Grb2 and Src are likely to be associated with ADAM 12 *in vivo*. More recently, further investigation into the interaction of ADAM 12 with

SH3 domains revealed that the cytoplasmic tail could bind to the SH3 domain of the p85α regulatory subunit of PI3K, both *in vitro* and *in vivo* (Kang et al., 2001). Three p85α binding sites in ADAM 12 involving PXXP motifs were identified, each of which is sufficient to mediate interaction with p85α *in vitro*, according to mutagenesis studies. Importantly, PI3K is essential for terminal differentiation of skeletal muscle cells. Transmembrane ADAM 12, by providing docking sites for p85α, could therefore play a role in the activation of PI3K by directly recruiting it to the plasma membrane. Since PI3K is critical for terminal differentiation of myoblasts and because expression of ADAM 12 is significantly up-regulated at the onset of myoblast differentiation, ADAM 12-mediated recruitment to the membrane may constitute one potential regulatory mechanism for PI3K during the differentiation process.

In conclusion, the ADAM protein family and, in particular ADAM 12, have been the focus of many studies over the past decade. The presence of these multiple domains in a given polypeptide is an intriguing finding and more work will be needed to understand the mechanism by which these proteins make use of their multiple, potential functions *in vivo*. In addition, understanding of the structural organisation of these domains will also shed light on the possible functions and regulation of this complex protein family.

3.3 Aims of this work

My work consisted in studying the extracellular domains of ADAM 12, and more particularly the disintegrin and cysteine-rich domains. ADAM 12 appeared to be an interesting protein for further study. Among the newly discovered molecules involved in muscle cell development, ADAM 12 was one of the rare

molecules to be an integral membrane protein and have no known function. Most of the potential functions of ADAM 12 are also achieved by other proteins (adhesion molecules, matrix metalloproteinases...) and it was of much interest to understand how a single protein may coordinate its multiple domains, so as to accomplish several functions. Notably, the disintegrin and cysteine-rich domains were of great interest because of the functions they may be involved in (namely adhesion and potentially fusion). No structural data is available so far and database searches also reveal no known folds in these two domains. On structural terms, the disintegrin and cysteine-rich domains of ADAM 12 are very particular in that together, they contain a large number of cysteine residues (34), which is rather unusual for proteins and thus, all the more interesting. To further understand the role of ADAM 12 in myoblast fusion, we therefore performed both structural and functional studies on ADAM 12.

Chapter 2: Structural Characterisation of hADAM 12

1. Expression of ADAM 12

1.1 Bacterial expression

The first system tested for expression of human ADAM 12 was *Escherichia coli*. A number of constructs were prepared to improve the level of expression and the solubility of the protein. In addition, a variety of tags were used for the purification of the fusion proteins. Table 2 is a summary of the constructs used for expression in *E. coli*.

	Domains	Fusion protein	Vector	Residues	Detectable expression
1	Dis and cys	N-terminal His-tag	pRSET (NheI/EcoRI)	417-707	No
2	Dis and cys	MBP	pMal-c2g (BamHI/HindIII)	417-707	No
3	Dis and cys	GST	pGAT (BamHI/HindIII)	417-707	Yes
4	Cys	N-terminal His-tag	pRSET (NheI/EcoRI)	513-707	Yes
5	Deleted Cys[T]	N-terminal His-tag	pRSET (NheI/EcoRI)	564-707	Yes
6	Dis and cys	C- terminal His-tag	pRSET (BamHI/EcoRI)	417-707	Yes
7	Dis and cys	N-terminal His-tag	pAB3 (AscI/SalI)	417-707	Yes

T This construct corresponds to a truncated form of the cysteine-rich domain (nucleotides 2000-2433).

Table 2: Summary of the constructs used for bacterial expression of ADAM 12.

Since the original construct (construct 1) corresponding to the disintegrin and cysteine-rich domains cloned with a N-terminal His-tag was not detectably expressed, several constructs of these same domains were prepared as fusion proteins, either with the maltose binding protein (MBP) or the glutathione-S-transferase (GST). Although no expression was detected with the MBP fusion protein, the expression level of the GST fusion protein (construct 3) was very good (figure 4A), upon induction with 1 mM isopropylthiogalactoside (IPTG). This fusion protein was expressed in the *E. coli* strain, BL21 (pUBS). However, the fusion protein was insoluble and was present as inclusion bodies in the bacteria. These inclusion bodies could be purified (as described in Materials and Methods) and solubilised in 8 M guanidine. This also served as an excellent protein purification step. The next step was, therefore, to refold the denatured protein. The disintegrin and cysteine-rich domains of ADAM 12 contain all together 34 cysteines, and therefore could potentially form 17 disulphide bonds. Two strategies were thus used: (i) the protein was diluted to a concentration of 2 μM in a buffer containing 6 M urea, 5 mM reduced glutathione and 0.5 mM oxidised glutathione, to allow the formation of disulphide bonds, and was then progressively dialysed against a buffer containing 2 M urea, and finally against a buffer without any urea, and (ii) the protein was added drop wise to a buffer containing 400 mM L-arginine HCl, 5 mM reduced glutathione and 0.5 mM oxidised glutathione, to a final protein concentration of 2 μM and was then dialysed to remove the L-arginine and glutathione. In the second case, the collected protein precipitated upon concentration. The first method yielded protein that could be concentrated to 0.5 mg per ml, although a considerable amount of protein was lost during concentration. Nevertheless, analysis of the

refolded protein by SDS-PAGE in the presence and absence of β-mercaptoethanol (which reduces the disulphide bonds) revealed that most of the protein was aggregated or formed intermolecular disulphide bonds, rather than intramolecular bonds (figure 4B).

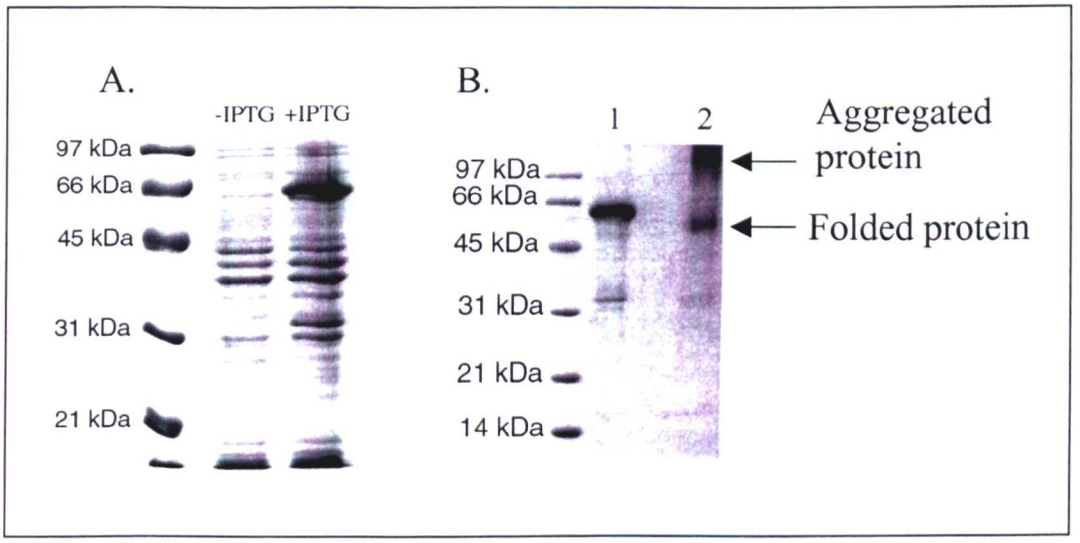

Figure 4: (A) SDS-PAGE analysis of the expression test of the disintegrin and cysteine-rich domains of ADAM 12 fused to the GST protein (construct 3). (B) Analysis of the purified GST fusion protein in the presence (lane 1) and absence of (lane 2) reducing agents.

Construct 6, corresponding to the disintegrin and cysteine-rich domains of ADAM 12 with a C-terminal His-tag, unlike the N-terminally tagged protein (construct 1) was well expressed but the refolding experiments yielded similar results as those obtained for the fusion proteins.

In view of these initial results, we prepared new constructs corresponding to shorter domains of ADAM 12, in the hope that these might fold more efficiently. A construct corresponding to the cysteine-rich domain alone (construct 4) and another corresponding to a truncated form of the cysteine-rich domain

(amino acids 564 to 707; construct 5) were cloned with a N-terminal His-tag. Gilpin *et al.* had previously expressed the latter construct for the purpose of producing a polyclonal antibody (Gilpin et al., 1998). Indeed, these shorter constructs did fold more efficiently. In particular, the truncated form of the cysteine-rich domain was expressed at a high level in BL21 (pUBS), purified as inclusion bodies and refolded according to the first method described above (figure 5). As opposed to the fusion proteins that appeared to form higher molecular weight aggregates after refolding, this smaller construct migrates at its appropriate molecular weight, both in the presence and absence of reducing agents.

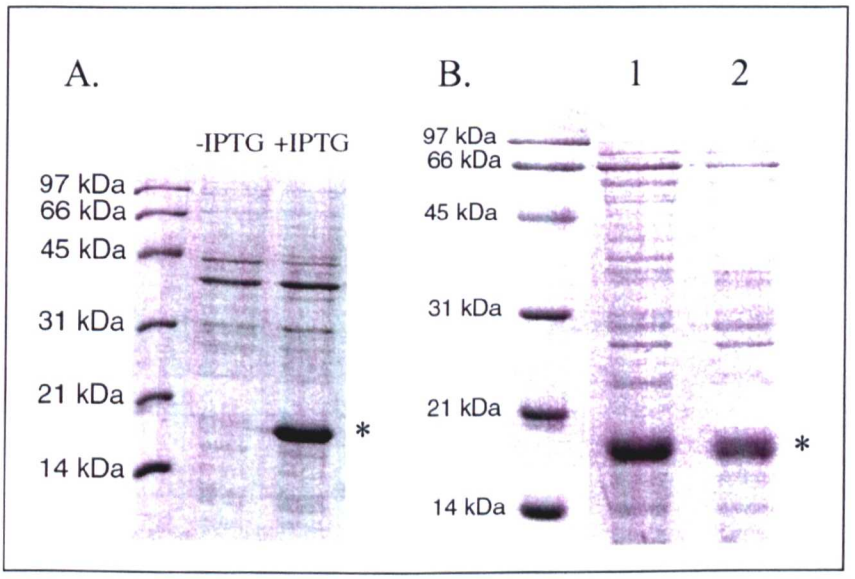

Figure 5: (A) SDS-PAGE analysis of the expression test of the truncated form of the cysteine-rich domain of ADAM 12 (construct 5). (B) Analysis of the purified and refolded, truncated construct (*) in the presence (lane 1) and absence of (lane 2) reducing agents.

Nonetheless, although the folding process appeared to be more efficient with this short construct, the protein was still very unstable and precipitated upon concentration or during further purification steps. Moreover, without an activity

test available for ADAM 12, it was impossible to know whether the protein was properly folded or not. After a final trial to express the disintegrin and cysteine-rich domains of ADAM 12 (construct 7) in the periplasmic space of *E. coli* (which allows for the formation of disulphide bonds), which yielded insoluble protein that was therefore not targeted to the periplasmic space, it was concluded that ADAM 12 might need further post-translational modifications, unavailable in bacterial expression systems, for its correct folding.

1.2 Expression in *Pichia pastoris*

Pichia pastoris has now become an attractive, alternative system available for efficient expression of extracellular proteins or single domains. As a yeast, *P. pastoris* is a unicellular microorganism that is easy to manipulate and culture. However, it is also a eukaryote and capable of many of the posttranslational modifications performed by higher eukaryotic cells, such as proteolytic processing, folding, disulphide bond formation, and glycosylation. Moreover, with *P. pastoris*, the recombinant proteins can either be expressed intracellularly or secreted into the medium. Because *P. pastoris* secretes only low levels of endogenous proteins and because its culture medium contains no added proteins, a secreted protein comprises the vast majority of the total protein in the medium. Thus, secretion serves as a major first step in purification, separating the foreign protein from the bulk of cellular proteins. Secretion of recombinant proteins is of particular interest for the expression of extracellular proteins, which possess a signal peptide, as it is the case for ADAM 12.

In the case of ADAM 12, the disintegrin and cysteine-rich domains were cloned into the pPICZα vector (EcoRI/XbaI sites; Invitrogen), which contains a *S.*

cerevisiae alpha-factor prepro signal sequence, a C-terminal *myc*-epitope and the *Sh ble* gene for selection of recombinant transformants with zeocin. Interestingly, selection of zeocin-resistant transformants at high concentrations of zeocin generates an enrichment in recombinant strains with multiple copies of the integrated vector. Thus, direct selection of zeocin hyperresistant transformants can be used to generate a population of multicopy clones that may ultimately result in an increase in the level of protein production. In the case of ADAM 12, however, colonies corresponding to transformants were only obtained on plates containing the lowest amount of zeocin (100 μg/ml). No colonies were obtained on plates containing 500, 1000 or 2000 μg/ml. These results suggested that these colonies were single-copy transformants. These results were obtained with three different strains of *P. pastoris*: KM71, X-33 and GS115. KM71 is a strain in which the chromosomal AOX1 gene is largely deleted and replaced with the *S. cerevisiae* ARG4 gene. As a result, this strain must rely on the much weaker AOX2 gene promoter for production of the enzyme alcohol oxidase (AOX), the enzyme required in the first step in the metabolism of methanol. Thus, KM17 strain grows very slowly on methanol (Mut^s phenotype), and this characteristic slow growth may improve the folding of recombinantly expressed proteins, although it also reduces the level of expression. GS115 is the most commonly used expression host, which is wild-type with regards to the AOX1 and AOX2 genes and grows on methanol at the wild-type rate (Mut^+ phenotype). Finally, the X-33 strain, which is isogenic to GS115, was derived as a strain ideally suited for expression in fermentation without supplemental histidine.

Since the colonies obtained were most likely single-copy transformants, the GS115 strain, which should display a higher expression level compared to the

strains X-33 and KM71, was used to test for expression. Several GS115 transformants were thus used to inoculate cultures and were induced to express the recombinant protein by transferring the yeast to methanol-containing medium. Aliquots of both cell extracts and cell supernatants were analysed by western blot at various time points (every 12 hours) after induction. Since the ADAM 12 gene had been cloned in frame with a C-terminal *myc*-epitope, an anti-*myc* antibody was used for detection of the recombinant protein. However, no detectable protein could be seen on the western blots (neither in the supernatants nor in the cells). A possible explanation for this is of course that multicopy transformants may be needed to detect the expression of the ADAM 12 disintegrin and cysteine-rich domains. Secreted proteins are commonly expressed at a lower level than are cytoplasmic proteins in *P. pastoris*. Based on available data, there is a 50-75% probability of expressing a given protein in *P. pastoris* at a reasonable level, and the biggest difficulty appears to be to generate initial success, *i.e.* expressing a protein at any level (Higgins and Cregg, 1998). In addition, degradation of proteins by proteases can be a serious problem in the *P. pastoris* system, and proteolysis studies on the disintegrin and cysteine-rich domains of ADAM 12 produced in *E. coli* revealed that these domains were very susceptible to protease digestion (data not shown).

1.3 Expression in Sf9 insect cells

Taking into consideration the expression difficulties we had faced with the bacterial and yeast systems, it was decided to try the baculovirus expression system in insect cells. The very strong polyhedrin promoter found in this virus has been used to produce large quantities of recombinant proteins in insect cells; these

proteins are correctly folded and glycosylated in a manner closely approximating that achieved in mammalian cells.

The cells used for this work were *Spodoptera frugiperda* cells (Sf9). Three different constructs (summarised in table 3) were prepared to allow both overexpression and purification of the disintegrin and cysteine-rich domains (construct 1) but also to ensure correct folding and targeting of the protein to the plasma membrane (constructs 2 and 3).

	Domains	Fusion protein	Vector	Residues	Detectable expression
1	Dis and cys	C-terminal His-tag	pMelBac (BamHI/EcoRI)	417-707	Yes
2	Dis, cys and TM	C-terminal His-tag	pMelBac (BamHI/SalI)	417-732	Yes
3	Dis, cys and TM	FLAG	pMelBac (BamHI/SalI)	417-732	Yes

Table 3: Summary of the constructs used for baculovirus expression of ADAM 12 in Sf9 insect cells.

All three constructs were cloned into pMelBac (Invitrogen), which contains a honeybee mellitin signal sequence for efficient secretion of the recombinant proteins into the cell culture supernatant. During transfection, recombination between homologous sequences in the viral DNA and the transfer vector supply the essential sequence needed for replication of recombinant baculovirus. With vectors such as pMelBac that contain only a 5′ portion of the *LacZ* gene, recombination occurs between the *LacZ* and ORF1629 sequences, forming blue, recombinant plaques on medium containing 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). These blue plaques were then isolated and added to the medium of cells to make the initial virus stocks, P1 viral stocks. All the successful viral stocks used were titered and presented a titre of $2x10^8$ pfu/ml. These original

P1 viral stocks (2 ml) were used to prepare a large-scale high-titre virus stock of one litre by infecting cells at a multiplicity of infection (MOI) of 0.5. A low MOI was used in this case to allow for exponential increase in virus titre. To infect Sf9 cells for protein expression, however, a MOI of 5 was used, to infect the cells synchronously to produce protein at the same time. 72 hours post-infection, the disintegrin and cysteine-rich domains of ADAM 12 could be detected by western blot in both the cells and the supernatant (figure 6A).

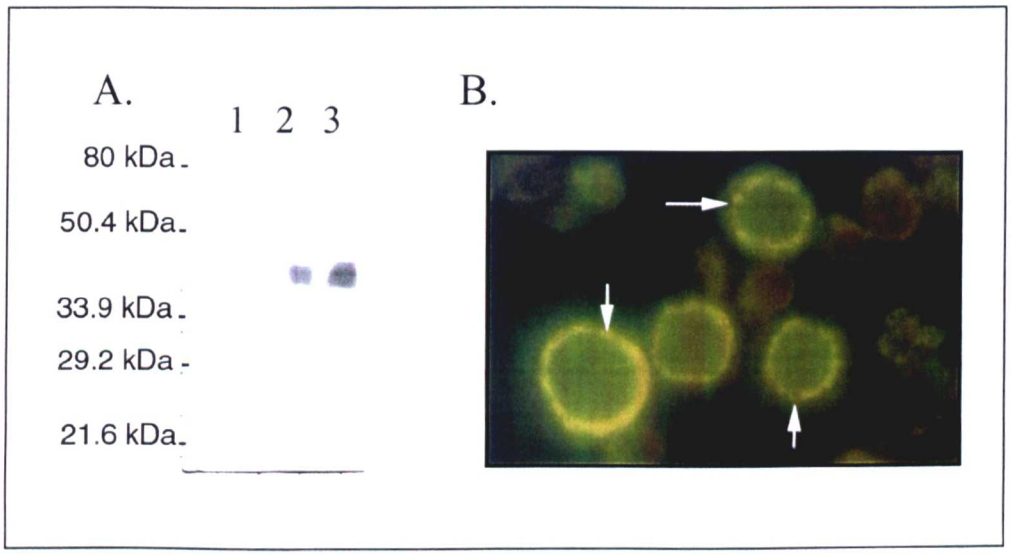

Figure 6: Expression of ADAM 12 in Sf9 insect cells. (A) Western blot analysis (using the anti-ADAM 12 antibody) of the expression of construct 1 in the supernatant of cells infected with a MOI of 1 (lane 1), 5 (lane 2) and 10 (lane 3). (B) Immunofluorescence staining (using anti-His antibody) of cells infected with recombinant baculovirus expressing construct 2.

From figure 6 we can see that the disintegrin and cysteine-rich domains of ADAM 12 (construct 1) are properly secreted into the cell culture supernatant and that construct 2, which contains an extra transmembrane domain, clearly localises to the plasma membrane of transfected cells. These initial results suggest that ADAM 12 is able to fold correctly and be properly targeted to its appropriate compartment in insect cells.

2. Characterisation of the disintegrin and cysteine-rich domains of ADAM 12

2.1 Large-scale production and purification of ADAM 12

For the purpose of large-scale expression and purification of the disintegrin and cysteine-rich domains of ADAM 12 (construct 1), two factors were crucial. The first was to obtain a single large high titre viral stock that could be used throughout the setting up of an efficient purification protocol. The second was to transfer the Sf9 insect cells, which are normally grown in medium containing 10% foetal bovine serum, to a serum-free medium to facilitate the purification of ADAM 12 from the cell culture supernatant. This was achieved by progressively passing the insect cells into medium containing less and less serum, over a period of two to three weeks. To obtain a large high titre viral stock, it was essential to amplify the stock directly from the initial 2 ml P1 viral stocks. If intermediate viral stocks were used for amplification, the expression was lost; this may be due to some DNA recombination, taking place between different viruses. But, finally a 1 litre stock of recombinant baculovirus was obtained at a titre of $2x10^8$ pfu/ml. This viral stock was sufficient to infect a total of 25 litres of Sf9 cells, grown in suspension in spinner flasks, at a MOI of 5.

72-hours post infection the cultured cells were pelleted and the supernatant was precipitated with 55% ammonium sulfate for 45 minutes at 4°C. The main purpose of this step was to reduce the total volume of protein before loading it onto an affinity column. In addition, the ammonium sulfate precipitation also served as an initial purification step, since not all the proteins in the cell culture supernatant were precipitated at 55% ammonium sulfate. The precipitated protein was then resuspended in one tenth of its original volume in a buffer containing 50

mM Tris-HCl pH 8.0 and 100 mM NaCl, before being loaded onto a chelating Ni-sepharose column, which binds the C-terminal His-tag of ADAM 12. After washing the column extensively the protein was eluted with an imidazole gradient (figure 7A). The usual protein yield at this step was 1 mg of protein per litre of cell culture supernatant. Finally, the pooled fractions from the affinity column were separated on a size exclusion column (Superdex 75) to remove the last few impurities and the imidazole (figure 7B).

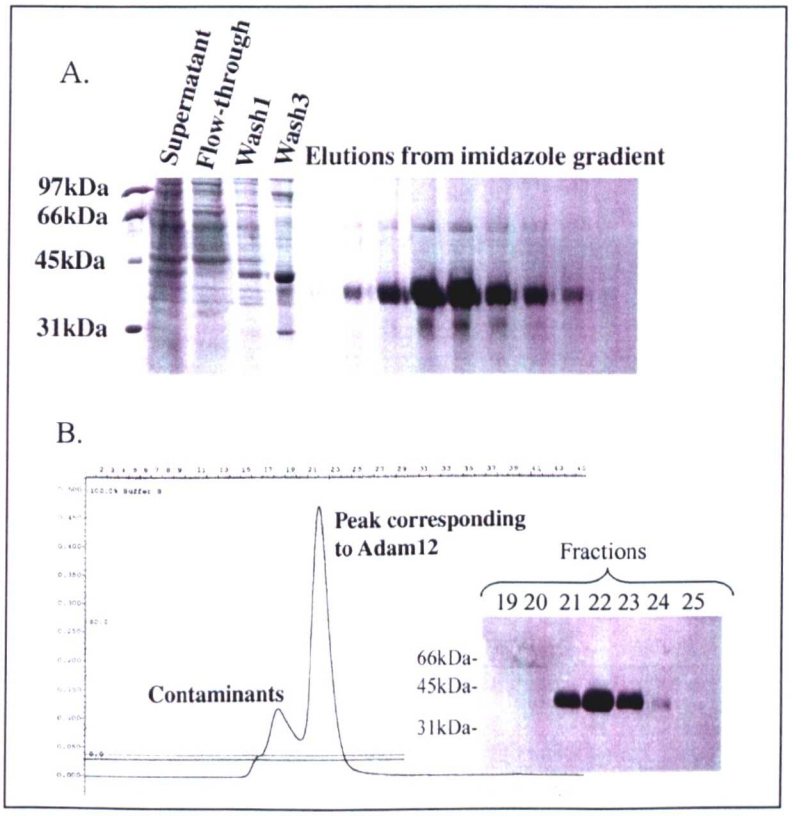

Figure 7: (A) SDS-PAGE analysis of the first purification step of ADAM 12 on a Ni-affinity column. (B) Chromatogram and SDS-PAGE analysis of the size exclusion chromatography step.

As we can see from figure 7B, the protein obtained after the size exclusion chromatography is more than 99% pure and elutes at 11 ml, which clearly corresponds to a monomer (calculated molecular weight is 31.9 kDa). The column

was calibrated with bovine serum albumin (67 kDa), ovalbumin (43 kDa) and chymotrypsinogen A (25 kDa) that elute at 9.5 ml, 10.5 ml and 12.5 ml respectively.

2.2 Biochemical characterisation of ADAM 12

Initial characterisation of the purified disintegrin and cysteine-rich domains showed that the protein migrates as a single band on a native gel, which separates proteins according to their charge (figure 8B). When separated on SDS-PAGE three bands could be distinguished in the absence of β-mercaptoethanol (figure 8A), suggesting that the protein preparation was not homogeneous, and that several disulphide bond patterns have formed during the folding of the protein in the insect cells.

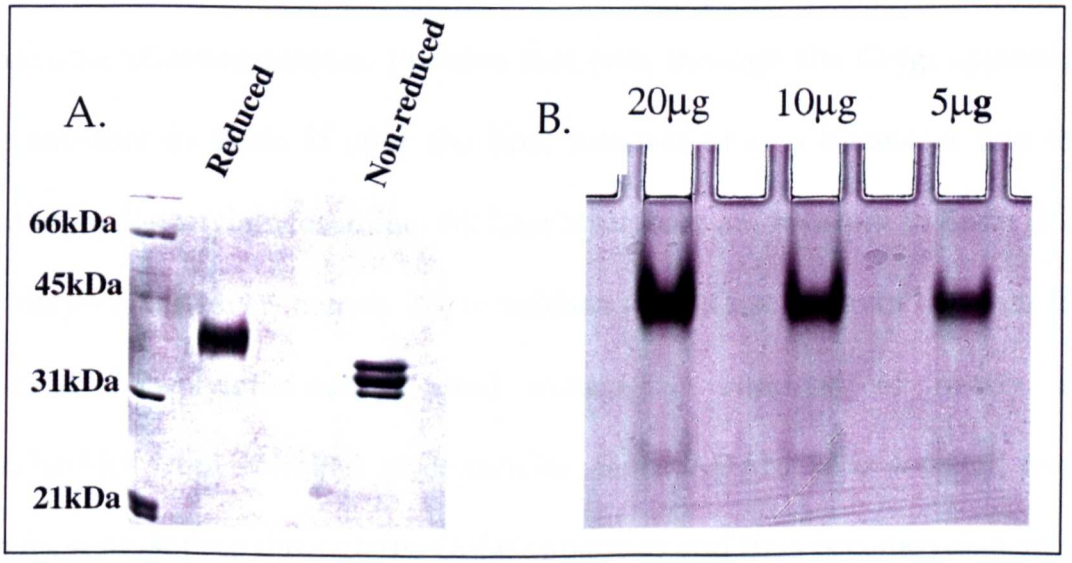

Figure 8: (A) SDS-PAGE analysis of the purified ADAM 12 in reducing and non-reducing conditions. (B) Analysis of different amounts of purified ADAM 12 on an 8% non-denaturing gel at pH 8.8.

Trials to separate the three distinct bands using a chromatofocusing column (Mono P) failed. A Mono P column consists of MonoBeads substituted with

positively charged amine groups. These columns greatly concentrate samples and resolve pI differences of 0.02 pH units. However, this was not sufficient to separate these three bands, which appear to have a very similar overall charge (a single band on a native gel).

From the SDS-PAGE analysis, the disintegrin and cysteine-rich domains of ADAM 12 migrate at a molecular weight between 35 and 40 kDa, whereas the calculated molecular weight of these two domains (total length of 297 amino acid residues) is 31.9 kDa. Interestingly, within these two domains there are two potential glycosylation sites. We therefore investigated to see whether the insect cell-produced protein had been posttranslationnally modified. To study the glycosylation state of ADAM 12 two enzymes were used. Endoglycosidase H (Endo H) cleaves the bond between the two *N*-acetylglucosamines in the core of high mannose oligosaccharides. Proteins that pass through the Golgi apparatus become resistant to Endo H after the final removal of two mannoses and the addition of an *N*-acetylglucosamine. All later structures are resistant to Endo H. A second enzyme, however, namely *N*-glycosidase F (PNGase) cleaves between the inner most *N*-acetylglucosamine and asparagine residues of nearly all oligosaccharides from *N*-linked glycoproteins. Since ADAM 12 is secreted from insect cells, it must pass through the Golgi apparatus and thus become resistant to Endo H cleavage. Indeed, this was the case. Treatment of both the native and denatured ADAM 12 with Endo H had no effect on the protein (data not shown), whereas treatment with both Endo H and PNGase resulted in a shorter form of the disintegrin and cysteine-rich domains, which migrated faster on SDS-PAGE (Figure 9).

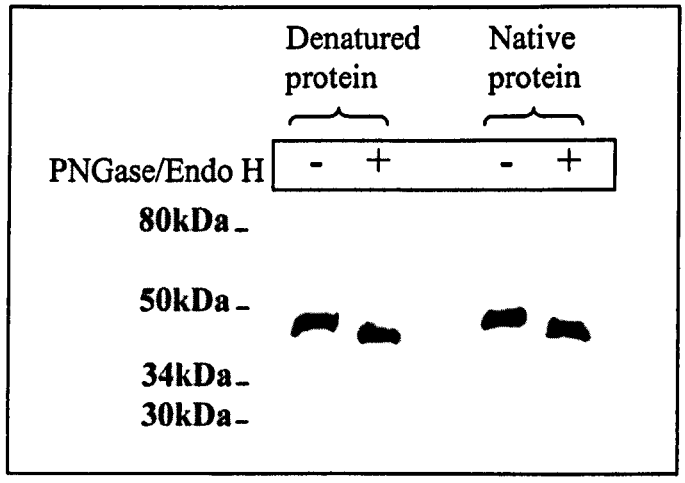

Figure 9: Western blot analysis (using the anti-ADAM 12 antibody) of the glycosylation state of ADAM 12 in the absence and presence of the two enzymes PNGase/Endo H.

The extent of glycosylation appeared to be very similar for the different forms of ADAM 12, since only a single band corresponding to ADAM 12 is obtained after deglycosylation. The fact that ADAM 12 is glycosylated in insect cells, further confirms the idea that these two domains must be folded, so as to present the two potential glycosylation sites at the surface of the protein.

Like fertilin-α, it was originally proposed that ADAM 12 contained a putative fusion peptide within its cysteine-rich domain. If this were the case, the fusion peptide of ADAM 12 should be able to insert into lipid membranes in a similar way to viral fusion peptides. To check this hypothesis, the supernatant of infected Sf9 cells expressing the disintegrin and cysteine-rich domains of ADAM 12 was incubated with liposomes containing 50% L-α-phosphatidyl-L-serine, 25% cholesterol and 25% phosphatidylcholine. The mix was then run through a sucrose gradient, and the fractions obtained were analysed by Western blot (figure 10). The liposomes were clearly visible in the upper fractions of the gradient (fractions

1-3), while the protein remained in the bottom, implying that ADAM 12 was not able to bind lipid membranes *in vitro*.

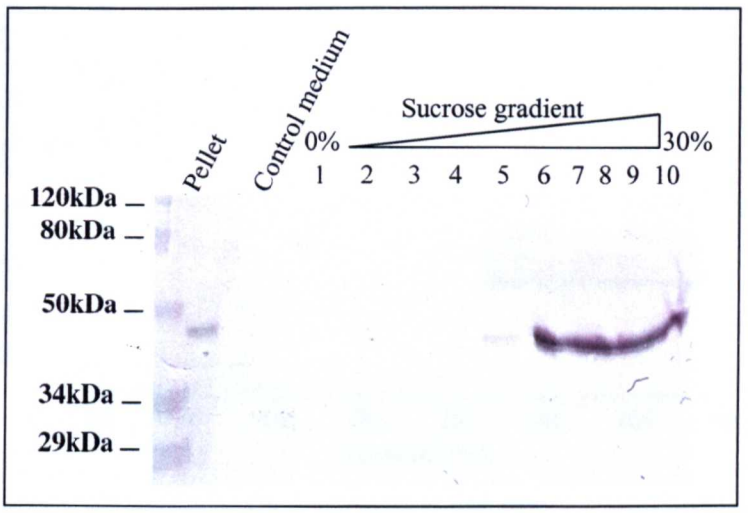

Figure 10: Western blot analysis, using the anti-ADAM 12 anti-serum, of the liposome floatation assay with the insect-cell produced ADAM 12.

2.3 Structural characterisation of ADAM 12

Circular dichroism (CD) measurements were carried out on the insect cell-purified disintegrin and cysteine-rich domains of ADAM 12. Interestingly, these initial measurements revealed that these two domains had very little secondary structure and the measured CD spectrum resembled that of a random coil conformation (figure 11). A randomly arranged polypeptide chain has a negative CD band centred around 200 nm and displays no positive bands (Johnson, 1990). Although it is difficult to obtain much structural information from this spectrum, this result suggests that these domains may be unstructured and consist of a series of loop regions, held together by the numerous (17) disulphide bonds, in a similar fashion to the EGF module. CD measurements were also carried out at increasing temperatures (up to 95°C) and no significant changes were observed in the spectrum.

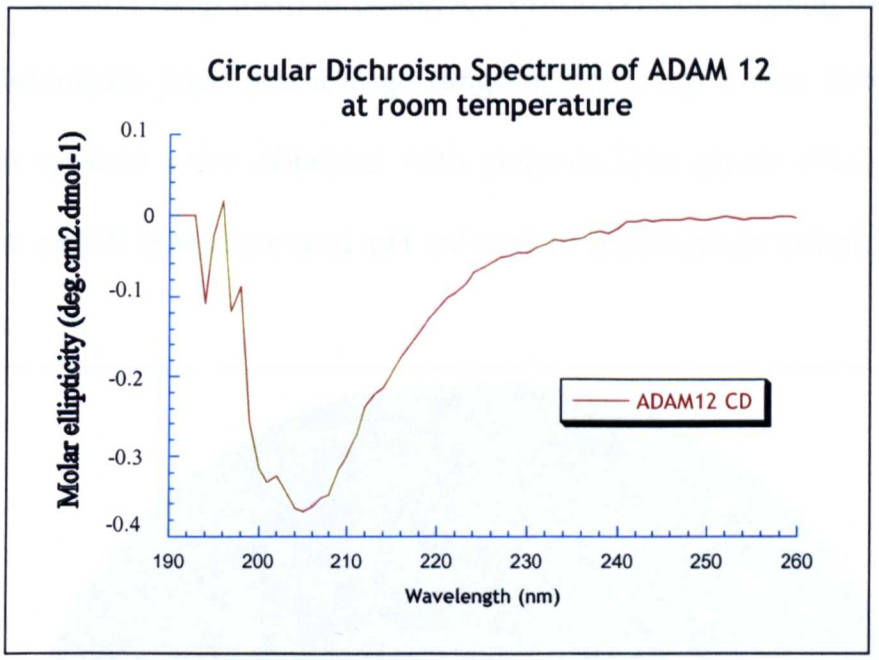

Figure 11: Circular dichroism spectrum of the disintegrin and cysteine-rich domains of ADAM 12 at room temperature.

In addition to circular dichroism, many crystallisation trials were set up in order to obtain a detailed structure of these two domains. For this purpose, the purified protein collected from the size exclusion column was concentrated to 8-10 mg/ml.

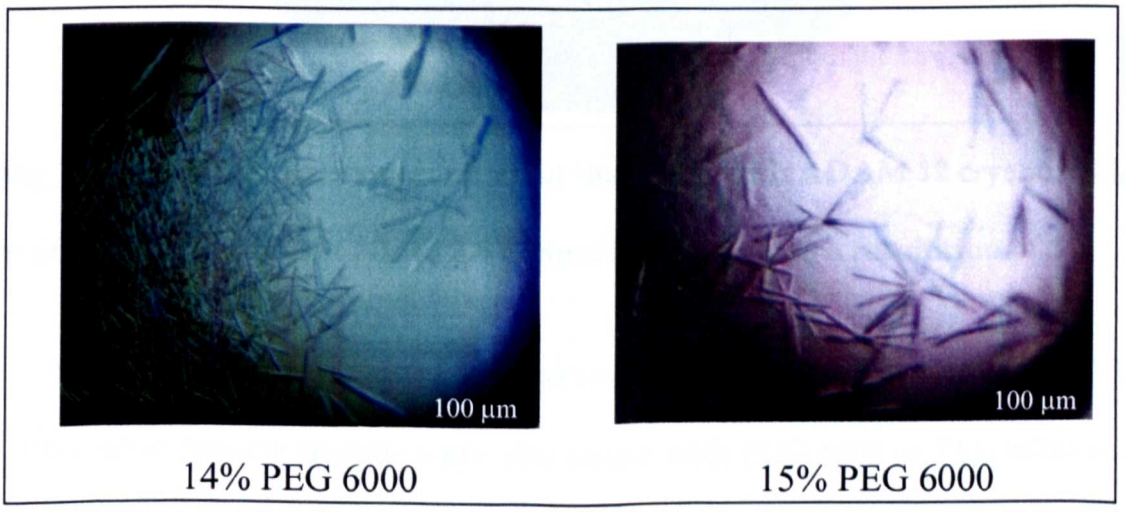

Figure 12: Examples of the needle-like crystals of the disintegrin and cysteine-rich domains of ADAM 12.

The hanging drop method (using a 1:1 ratio (v/v) of protein to precipitant) was predominantly used and a large range of precipitants was tested. The first needle-like crystals were obtained with polyethylene glycol (PEG) 6000 as a precipitant, in 100 mM citric acid pH 4.0 and 10 mM barium chloride dihydrate (figure 12).

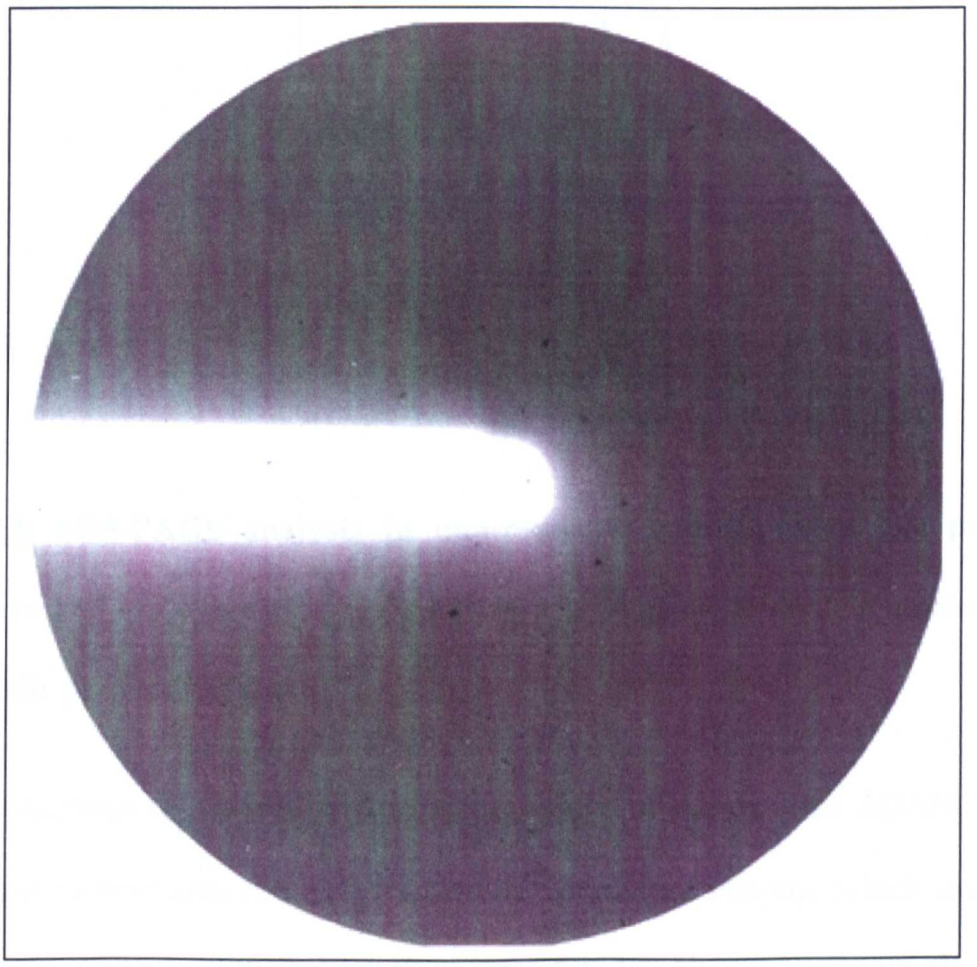

Figure 13: Diffraction pattern obtained at the ESRF with ADAM 12 crystals. The edge of the image corresponds to 2.6 Å. Spots are visible to a resolution of 5 Å.

Needles could be obtained in a concentration of PEG varying from 14 to 18% PEG 6000. Similar crystals were also found with PEG 4000 or PEG 8000 as a precipitant. However, further crystallisation trials and attempts to improve the quality of these crystals failed. These needle-like crystals, unfortunately, only diffracted to a low resolution (figure 13) at the European Synchrotron Radiation Facility (ESRF). There are several possible explanations for the poor quality of

these crystals. An SDS-PAGE analysis, under non-reducing conditions, of the protein contained in the crystal, revealed that all three forms of the disintegrin and cysteine-rich domains were present in the needles (figure 14). This heterogeneity within the crystals may well explain why it was difficult to grow larger and more regular crystals.

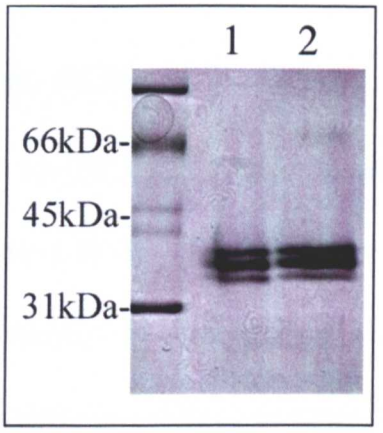

Figure 14: SDS-PAGE analysis in non-reducing conditions of the purified disintegrin and cysteine-rich domains (lane 1) and of the crystals obtained from this protein preparation (lane 2).

In addition to these different conformations of the purified ADAM 12, the crystallised protein also contained a flexible C-terminal His-tag, which may also influence the packing of the protein in the crystal. A new construct was, therefore, prepared that included a Factor Xa protease cleavage site in between the His-tag and the C-terminus of the cysteine-rich domain, to allow the removal of the His-tag after the purification. Although new recombinant baculovirus stocks were obtained that expressed this new construct in insect cells, the expression level was significantly lower and the protein obtained after the two-step purification was largely contaminated with other proteins. A sufficient amount of pure protein could not be obtained to set up new crystallisation trials.

To improve the crystallisation of the disintegrin and cysteine-rich domains of ADAM 12, a search for a possible protein partner was carried out, that may be able to stabilise the domains further and thus increase their ability to crystallise.

3. Heterodimerisation of ADAM 12 and ADAM 19

3.1 Cloning of human ADAM 19

When the mouse gene for ADAM 12 (meltrin α) was initially isolated, two other cDNAs encoding members of the ADAM family were also isolated: meltrin β and γ. By homology to fertilin α and β that form a heterodimer, it was hypothesised that ADAM 12 may also need a partner for its role *in vivo*. From the tissue distribution pattern of meltrin β and γ, it appears that meltrin β is a better candidate as a partner for ADAM 12, since they display a very similar expression pattern, whereas meltrin γ is ubiquitously expressed.

Since human meltrin β (ADAM 19) had not yet been cloned, the cDNA encoding ADAM 19 was isolated from a cDNA library. Using the mouse meltrin β as a query sequence for a BLAST® (Basic Local Alignment Search Tool) search, a human expressed sequence tag (EST) was isolated (300 base pairs long) which displayed 84% sequence identity with the mouse cDNA sequence. The EST sequence corresponded to the end of the disintegrin domain and the beginning of the cysteine-rich domain. Based on this EST sequence, oligonucleotides corresponding to the 5′ and 3′ end of the EST were used to amplify by PCR the 300 base pair (bp) fragment from a human skeletal muscle cDNA library. In order to obtain more sequence information so as to amplify the full ADAM 19 cDNA, the mouse meltrin β sequence was used to search the newly available human DNA

sequence database (the High Throughput Genomic Sequences or HTGS database). This database was created to accommodate a growing need to make 'unfinished' genomic sequence data available. This search was very successful and enabled the retrieval of most of the human ADAM 19 sequence (figure 15).

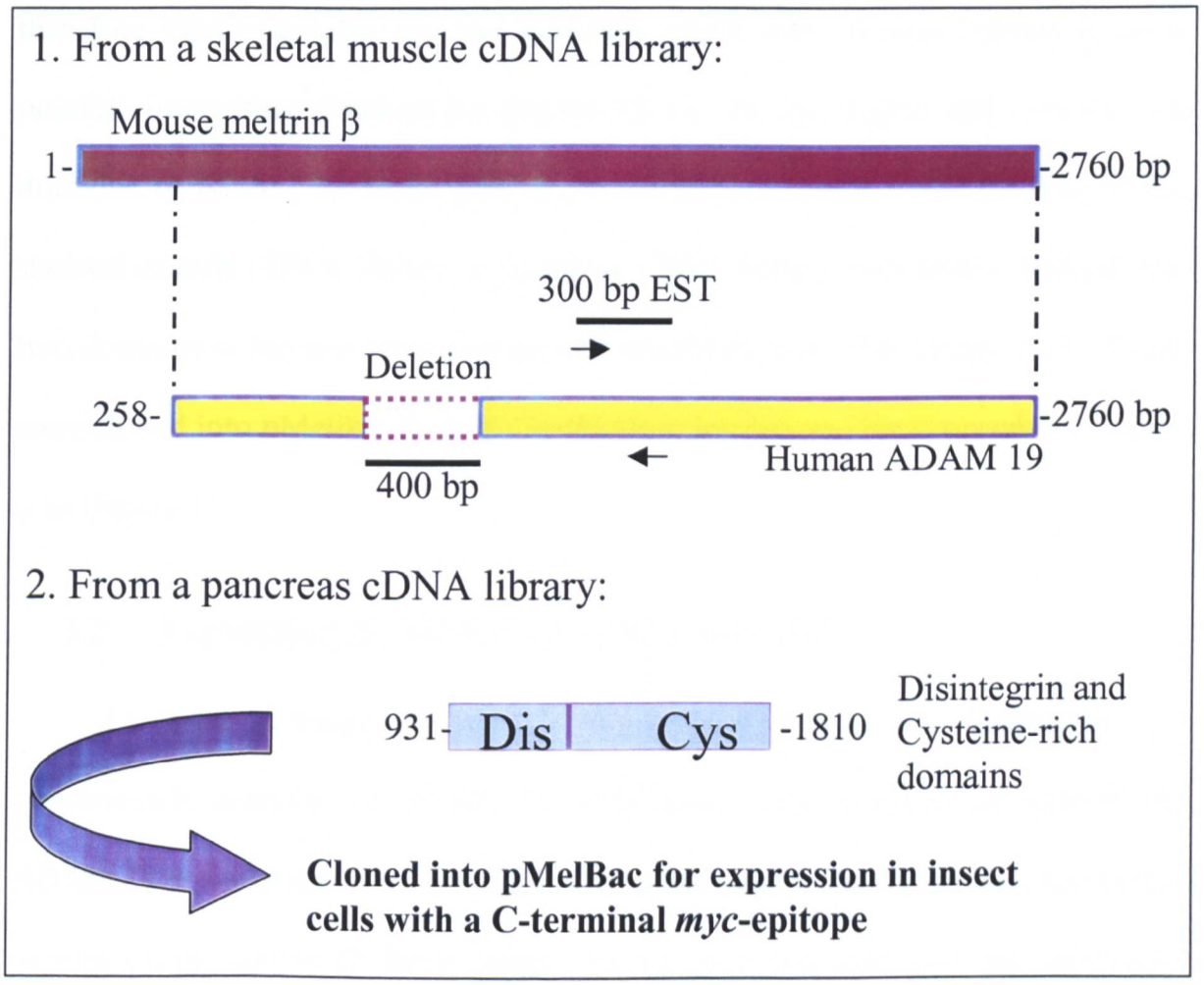

Figure 15: Schematic diagram of the cloning technique used to amplify the human ADAM 19 cDNA (in yellow) from a skeletal muscle cDNA library. In a second step the disintegrin and cysteine-rich domains (in blue) were amplified from a pancreas cDNA library.

However, the 5′ end (257 bp) was still missing. Comparison of the obtained human ADAM 19 sequence with that of mouse showed that the genes were very highly conserved, since they displayed 84% sequence identity. Appropriate oligonucleotides were then used to amplify by PCR the full-length ADAM 19 from

the human skeletal muscle cDNA library. Sequencing of the obtained fragments revealed, nevertheless, that approximately 400 bp were missing from the end of the metalloprotease domain and the beginning of the disintegrin domain (figure 15). Further trials to obtain the full-length clone of human ADAM 19 failed. We therefore tried to amplify the two domains, which were of most interest to us as possible interacting domains for ADAM 12, i.e. the disintegrin and cysteine-rich domains of ADAM 19. Since part of the disintegrin domain was missing in the skeletal muscle cDNA library, a pancreas cDNA library was tested. Indeed, the two domains of interest were successfully amplified from this library by PCR and were cloned into pMelBac (BamHI/EcoRI sites; Invitrogen) for expression in insect cells (figure 15).

3.2 Expression of ADAM 19 in Sf9 insect cells

In order to study the possible interactions between the disintegrin and cysteine-rich domains of ADAM 12 with these same domains of ADAM 19, ADAM 19 was expressed in insect cells, since this system had allowed a successful expression of ADAM 12. Recombinant baculovirus was prepared that efficiently expressed our gene of interest with a C-terminal *myc*-epitope. Since the baculovirus produced ADAM 12 possesses a His-tag at its C-terminus, we set up a pull-down assay (figure 16), where ADAM 12 was first bound to a Ni-affinity column and in a second step, the cell culture supernatant containing ADAM 19 was loaded on the column. After extensive washing of the column, the proteins were eluted. The eluted proteins were analysed by SDS-PAGE followed by a western blot using an anti-His antibody to detect ADAM 12 and an anti-*myc* antibody for detection of ADAM 19.

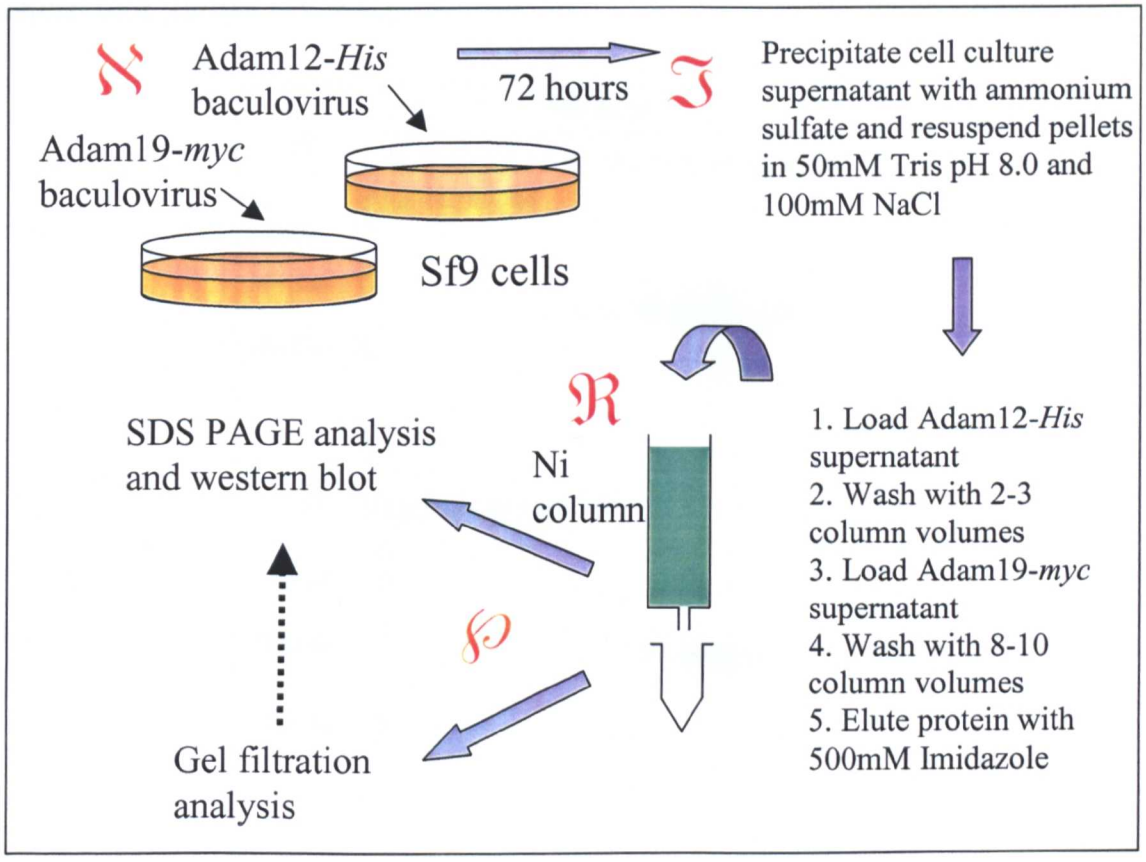

Figure 16: Schematic diagram of the steps involved in the co-purification of the disintegrin and cysteine-rich domains of ADAM 12 and ADAM 19 from Sf9 insect cells.

Both proteins were present in the eluate (data not shown). To further confirm the interaction of ADAM 12 with ADAM 19 and to ensure that the co-elution from the Ni-column was not a result of non-specific aggregation, two techniques were used: gel filtration and immunoprecipitation.

3.3 Evidence for heterodimerisation

The eluted proteins collected from the Ni-affinity column were first subjected to a size exclusion column (Superdex 200). As a control, purified ADAM 12 alone was also separated on the Superdex 200 (figure 17A). The fractions

collected from the size exclusion column were analysed by SDS-PAGE and western blot (figure 17B & C).

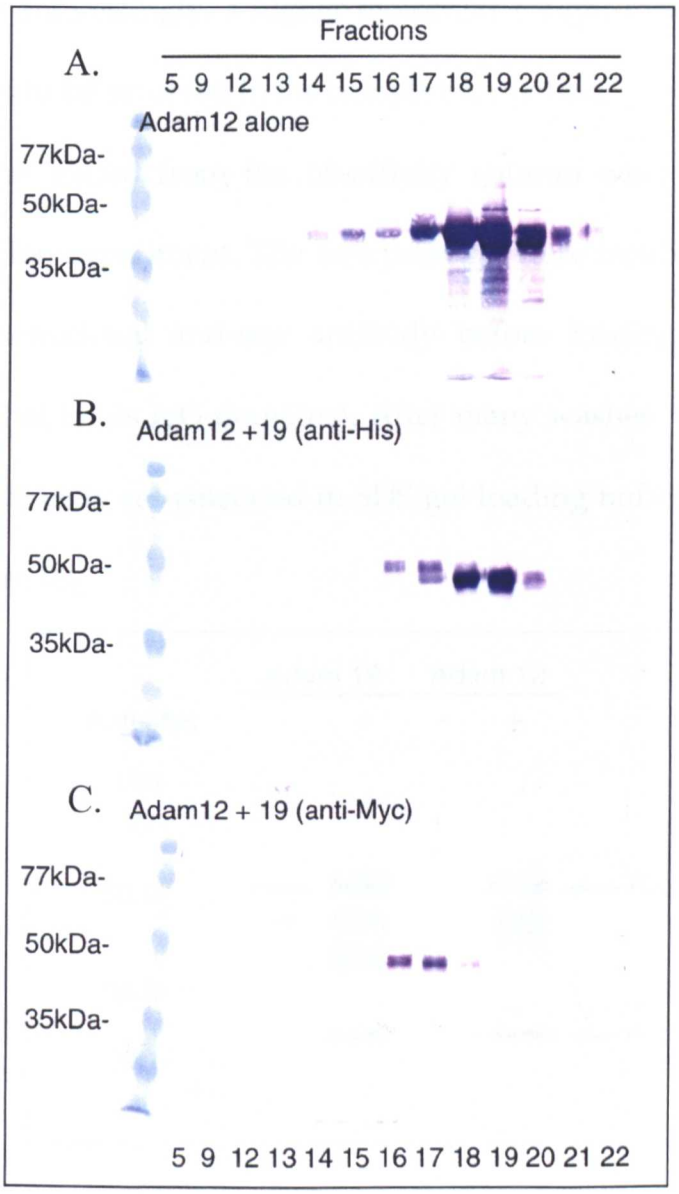

Figure 17: Western blot analysis of the fractions collected from the size exclusion chromatography of ADAM 12 alone (A) and of both ADAM 12 and ADAM 19 (B & C). ADAM 12 was detected using the anti-His antibody (A & B) and ADAM 19 was detected using the anti-*myc* antibody (C).

The gel filtration analysis shows that ADAM 12 alone elutes mainly in fractions 18 and 19, whereas ADAM 12 and ADAM 19 when purified together both elute in fractions 16 and 17. An excess of ADAM 12 also elutes in fractions 18

and 19. The fact that they both elute in the same fractions and also earlier than does ADAM 12 alone, strongly suggests that they form a higher molecular weight complex together. Interestingly, a higher molecular weight form (upper band) of ADAM 12 appears to be involved in the complex formation.

The proteins eluted from the Ni-affinity column were also used for an immunoprecipitation experiment. The two proteins were incubated for one hour with a mouse monoclonal anti-*myc* antibody before loading on a protein A sepharose resin (that binds IgG domains). After many washes, the protein bound to the beads was directly resuspended in SDS gel loading buffer and analysed by western blot (figure 18).

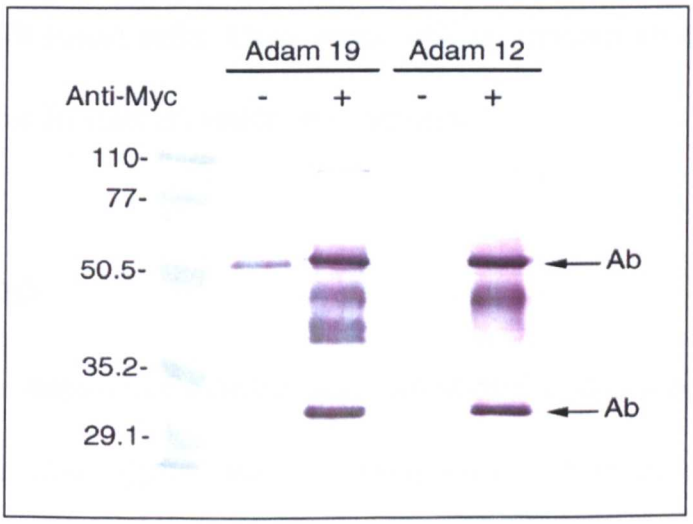

Figure 18: Western blot analysis of the immunoprecipitation experiment in the presence (+) and absence (-) of the anti-*myc* antibody. The same fractions were analysed either with the anti-*myc* antibody in the case of ADAM 19 or with the anti-His antibody for ADAM 12. The upper and lower bands (Ab) correspond to the added antibody used for the immunoprecipitation reaction.

Here again, it is clear that ADAM 12 is specifically co-immunoprecipitated with ADAM 19. No ADAM 12 could be detected in the absence of the anti-*myc* antibody, suggesting that the presence of ADAM 12 is not due to a non-specific

binding to the protein A sepharose resin. Interestingly in the immunoprecipitation experiment, three different forms of ADAM 19 and only a single, higher molecular weight band of ADAM 12 were precipitated. In the gel filtration analysis, the fractions 16 and 17 only contain the higher molecular weight forms of ADAM 12 and ADAM 19. These results suggest that higher molecular weight forms of both ADAM 12 and ADAM 19 are involved in the dimer formation. These higher molecular weight forms may result from extensive glycosylation of both ADAM 12 and ADAM 19.

Preliminary work has been done in order to co-express these two domains of ADAM 12 and ADAM 19 in a single expression vector, pFastBacDual (Life Technologies), in Sf9 insect cells. More work will be needed to obtain sufficient amount of protein for further crystallisation studies.

4. Discussion

ADAM 12 possesses five extracellular domains: the signal peptide, a pro-, a metalloprotease, a disintegrin and a cysteine-rich domain. No structural information is yet available for any of these domains. Although the atomic structure of the metalloprotease domain of ADAM 17, a distant member of the ADAM family, has been solved (Maskos et al., 1998), for the other domains there is still no further knowledge. It was therefore of much interest to try to gain further insight into the structure of the disintegrin and cysteine-rich domains of ADAM 12. Not only would this information be useful for our understanding of the structure-function relationship in the case of ADAM 12, it would also be very relevant for most of the ADAM family members.

In the present study, several expression systems were tested, each system having its advantages and drawbacks. For ADAM 12, as for most eukaryotic proteins, the need for posttranslational modifications, to ensure the proper folding of the protein, rules out the possibility of using bacteria as an efficient expression system. Moreover, in the case of the disintegrin and cysteine-rich domains of ADAM 12, the large number of cysteine residues present in the protein was also an extra difficulty in finding an appropriate expression system. Although finally a significant amount of pure ADAM 12 could be obtained with the baculovirus expression system in insect cells, high quality crystals could not be grown in order to solve the structure of the protein to atomic resolution. For crystallisation purposes, the protein must not only be very pure, it is also important that it be homogeneous in solution for correct packing in the crystal. The fact that three forms of the purified domains could be separated by SDS-PAGE under non-reducing conditions, strongly suggested that the protein preparation was heterogeneous. This may well explain why the crystals were very small, and grew principally in one dimension.

Calvete *et al.* determined the disulphide bond pattern of Catrocollastatin C, a disintegrin and cysteine-rich protein isolated from *Crotalus atrox* (Calvete et al., 2000), by N-terminal sequencing and mass spectrometry. A large number of zinc metalloproteinases of varying molecular weights and biological functions has been isolated from crotalid and viperid venoms. Over the past few years, structural studies on these proteinases have suggested their organisation into four classes, PI-IV (Jia et al., 1996). The disintegrin and cysteine-rich domains of SVMPs (class III) and ADAM molecules are characterised by the presence of 15 strictly conserved cysteine residues in the disintegrin domain and 13 strictly conserved

cysteines in the cysteine-rich domain, ordered in a distinctive sequence pattern. In addition the ADAM molecules have an extra EGF-like module (6 cysteines). Despite having conserved cysteine residues, snake venom disintegrins exhibit three different disulphide bond patterns (Calvete et al., 1997): S-S (I), S-S (II) and S-S (III). Catrocollastatin C has two extra cysteine residues (13 and 16) that form an additional disulphide bond that is characteristically found in the disintegrin-like domains of cellular metalloproteinases, such as the ADAM proteins. These two cysteines are missing in snake venom disintegrins, which lack a C-terminal cysteine-rich domain. Figure 19 shows the amino acid sequence alignment of Catrocollastatin C, the haemorrhagic protein HR1B (a class III SVMP) and ADAM 12. In figure 19, most of the predicted disulphide bonds are based on the experimentally determined bonds for Catrocollastatin C. Notably all the disulphide bonds in the cysteine-rich domain were identified using N-terminal sequencing and mass spectrometry. However, for the disintegrin domain, the full disulphide bond pattern was not established. The bonds between cysteines 1-4, 2-7, 12-15 and 13-16 were experimentally determined, while the other bonds were predicted on the basis that Catrocollastatin C most probably falls into S-S pattern II (which has been defined), and that all known class III SVMPs contain 8 disulphide bonds within their disintegrin domain (Calvete et al., 2000). In addition, compared to SVMPs and Catrocollastatin C, ADAM proteins have an extra EGF module at the C-terminal end of the cysteine-rich domain. The disulphide bonds within this module were predicted according to the defined EGF pattern: C1-C3, C2-C4 and C5-C6 (Abe et al., 1998). Moreover, alignment of the cysteine-rich domains of ADAM 12 with those of class III SVMPs and Catrocollastatin C reveals that the latter are lacking about 25 residues that have

been proposed to be a putative fusion peptide. This sequence is located between cysteines 24 and 25, which are linked by a disulphide bond (figure 19); thus this sequence most probably forms an accessible loop region.

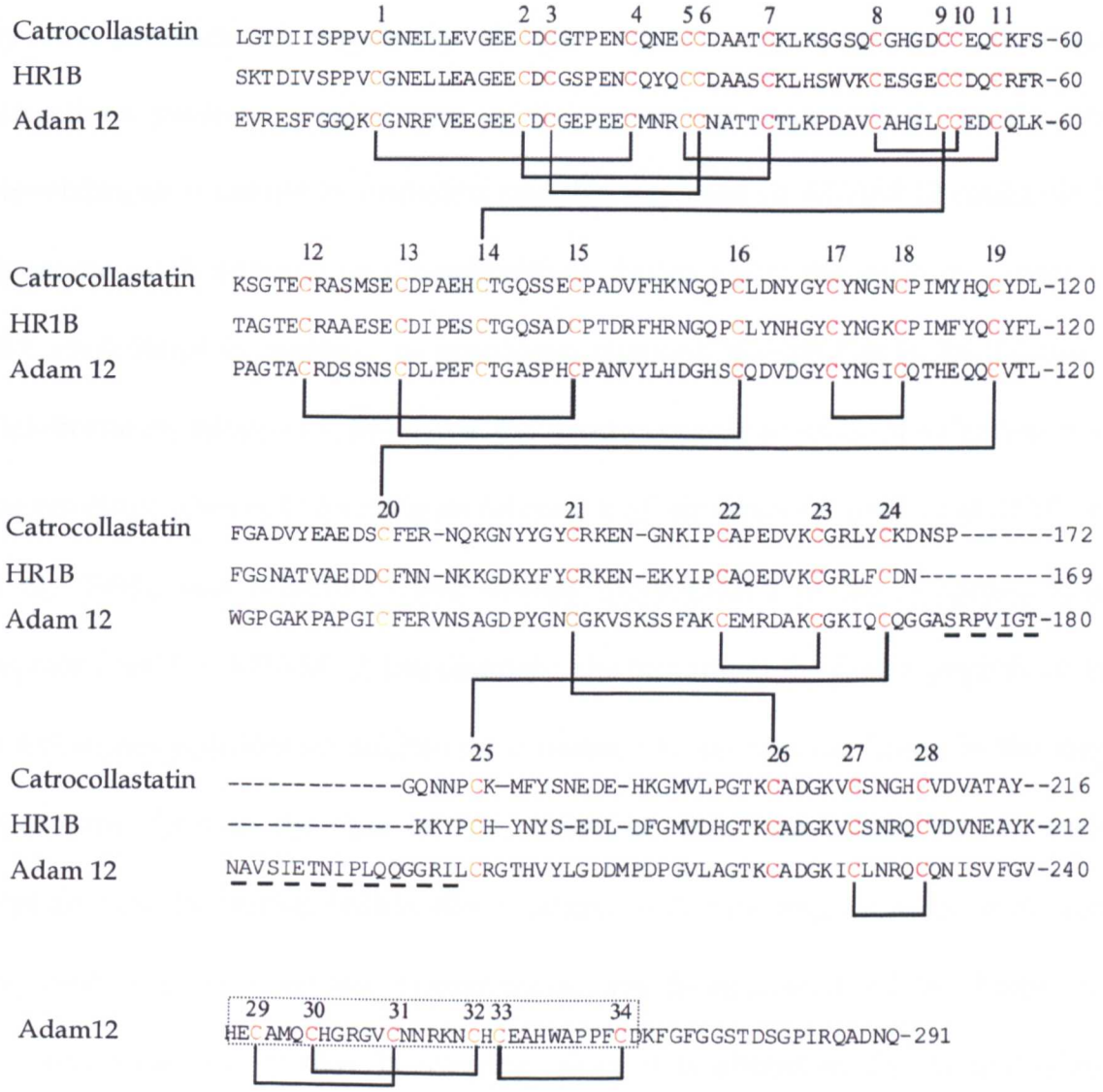

Figure 19: Comparison of the amino acid sequence of the disintegrin and cysteine-rich domains of Catrocollastatin-C, HR1B (a class III SVMP) and human ADAM 12. The possible disulphide bond pattern is shown as black lines. A dashed line shows the putative fusion peptide of ADAM 12, and the EGF module of ADAM 12 is boxed.

However, no ADAM molecule has been shown to possess bona fide membrane fusion activity and the finding that the cysteine-rich domain of ADAM

12 is unable to bind to liposomes *in vitro* suggests that this 25-residue stretch is most probably not a fusion peptide. So far there is evidence that a peptide representing the putative fusion domain of fertilin induces fusion of large unilamellar vesicles (Martin and Ruysschaert, 1997). This binding was enhanced by the presence of negatively charged lipids, suggesting that electrostatic attractions played a crucial role in the binding process (Martin et al., 1998). Nevertheless, it cannot be excluded that this sequence of ADAM 12 could bind to liposomes with different lipid compositions. In our assay, the liposomes contained 25% cholesterol in addition to negatively charged phospholipids. In the case of Tick-borne encephalitis (TBE) virus, the fusion peptide is believed to be internal to the envelope protein E, forming an internal loop structure (Allison et al., 2001; Rey et al., 1995), and therefore more similar topologically to the proposed fusion peptide found in ADAM 12. Interestingly, the binding of the fusion peptide of TBE to liposomes requires no additional proteins and no specific lipids in the target membrane. Like in the case of TBE, however, this loop containing the fusion peptide may be buried within the structure and may require a conformational change in order to insert into a target membrane. In all cases it will be of interest to find out what role it may be playing, since it is absent in the closely related SVMPs.

Cysteine 16 has been considered as the first cysteine of the cysteine-rich domain, while the experimental results obtained by Calvete *et al.* clearly demonstrate that cysteine 16 forms a disulphide bond with cysteine 13, implying that the domain borders may need to be revised. Cysteine 16 would then be the last cysteine of the disintegrin domain, and the molecules would thus form two separate domains, with no interdomain bonds (figure 20). This would also result

in an even number of cysteines in each domain, as opposed to the 15 and 13 cysteines originally described in the disintegrin and cysteine-rich domains respectively (Gilpin et al., 1998).

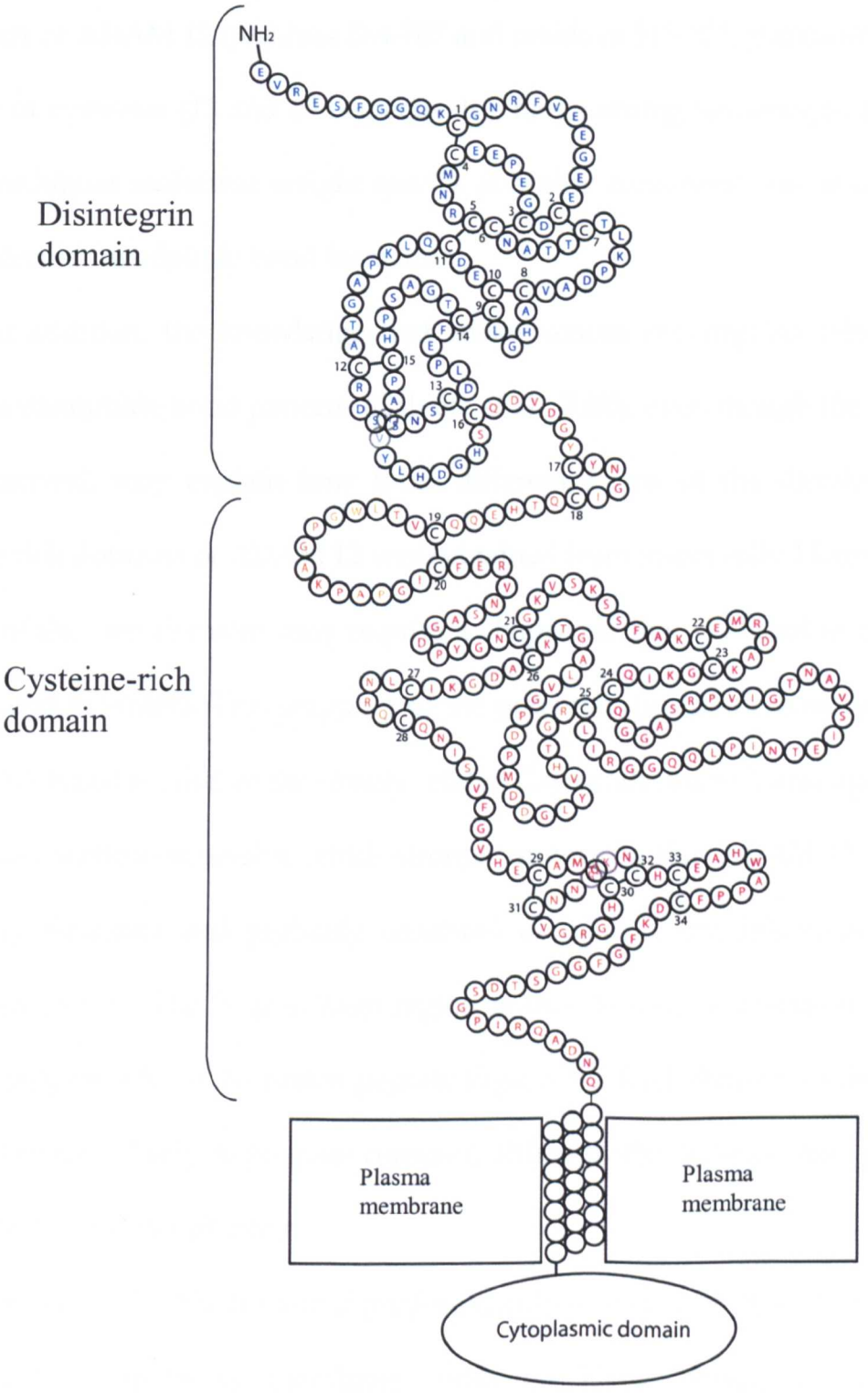

Figure 20: Model of the disintegrin and cysteine-rich domains of human ADAM 12 based on the deduced disulphide bond pattern. The amino acid residues of

the disintegrin and cysteine-rich domains are shown in blue and red respectively.

This new domain definition may also explain why the bacterially expressed constructs of ADAM 12 (residues 564-707 and residues 513-707) containing an odd number of cysteines (13 and 19 respectively), had a strong tendency to aggregate and form higher molecular weight species at higher concentrations, as a result of intermolecular disulphide bond formation.

In addition, the knowledge that snake venom disintegrins exhibit three different disulphide bond patterns (Calvete et al., 2000), even though the cysteines are conserved, may explain how three different forms of the disintegrin and cysteine-rich domains of ADAM 12 were obtained from insect cells. Homogeneous folding of the two domains may require a chaperone that is absent in the insect cell expression system. This mapping of the potential disulphide bond pattern of ADAM 12, based on that of the closely related Catrocollastatin C, also agrees with the circular dichroism results, which strongly suggested that ADAM 12 had little secondary structure and probably consisted of many loops interconnected by disulphide bonds. The longest loop region is this 26-residue insertion that was initially proposed to be the fusion peptide (figure 20). Each domain of the protein is therefore most likely to be quite compact, although the domains may be rather flexible relative to one another.

Disulphide bonds are found predominantly in secretory proteins and in the extracellular domains of membrane proteins. These bonds are important stabilising forces in the tertiary structure of these proteins. Site-specific mutagenesis studies have shown that proteins can be strengthened to resist

thermal denaturation by the introduction of disulphide bonds. A suitably positioned disulphide bond can raise the melting temperature of a given protein by 4°C or more (Matsumura et al., 1989).

Further stability may be achieved by expressing the full extracellular part of ADAM 12, rather than only two domains. It may be that the presence of the pro- and metalloprotease domains stabilises the overall tertiary structure of the extracellular domains of ADAM 12. In addition, the finding that the disintegrin and cysteine-rich domains of ADAM 12 can interact specifically with the corresponding domains of ADAM 19, also suggests that the contacts involved in the heterodimerisation may stabilise the domains. Finally, the use of nuclear magnetic resonance (NMR), as a tool for structure determination of proteins in solution, may also help to obtain further structural information concerning the disintegrin and cysteine-rich domains of ADAM 12. In this case, the limiting factor could be the size of the protein (35-40 kDa).

From a structural point of view the disintegrin and cysteine-rich domains of ADAM 12 and ADAM 19 must be very similar, since the protein sequence is 54% identical (figure 21). Notably, all the cysteines are conserved. It is also interesting to remark that the disintegrin and cysteine-rich domains of ADAM 12 are monomeric in solution and unable to form a homodimer. Thus, the interactions involved in the heterodimer formation must involve ADAM 12 and ADAM 19 specific residues, in addition to the possible requirement for the posttranslationnally added oligosaccharides. Within the disintegrin and cysteine-rich domains two putative glycosylation sites can be found in ADAM 12 and three in ADAM 19, which may explain why several bands of increasing molecular

weight can be distinguished by SDS-PAGE. Glycosylation is heterogeneous in insect cells, which explains the presence of these multiple bands.

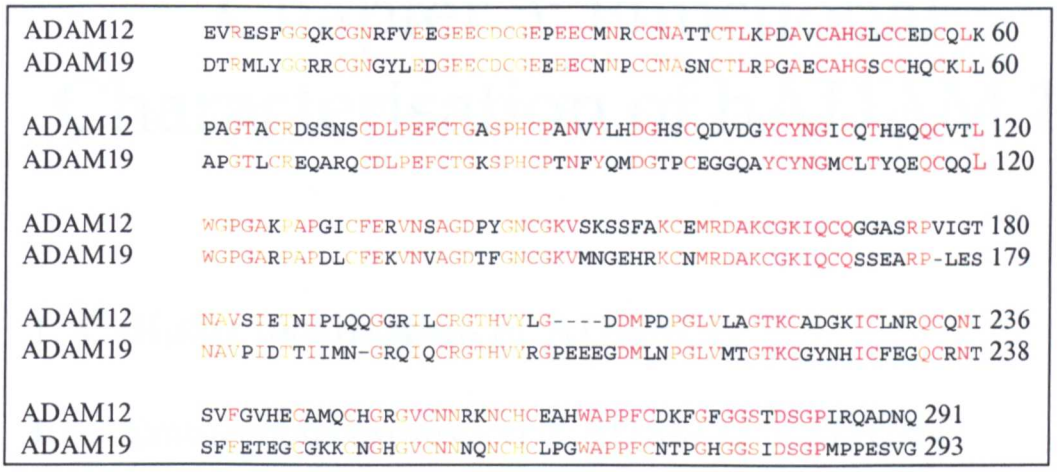

Figure 21: Amino acid residue alignment of the disintegrin and cysteine-rich domains of human ADAM 12 and ADAM 19 proteins. The identical residues are shown in red.

These studies clearly showed that the disintegrin and cysteine-rich domains of ADAM 19 could specifically interact, at least *in vitro*, with the corresponding domains of ADAM 12. Further studies will be needed to determine the functional significance of this interaction *in vivo*. Possible functional roles for ADAM 12 have been hypothesised based on its expression pattern during myogenesis, its tissue distribution and its multiple domains. However, there is still very little data available concerning an actual role of ADAM 12 in the developmental process of myoblasts. A more functional approach to the study of ADAM 12 was therefore also carried out.

Chapter 3: Functional Characterisation of hADAM 12

1. Cellular localisation of ADAM 12

1.1 Endogenous expression of ADAM 12 in RD cells

RD cells are human embryonal rhabdomyosarcoma cells (muscle tumour cells) that were obtained from the American Type Culture Collection (ATCC #CCL 136). The RD cell line is the first human rhabdomyosarcoma cell line characterised and studied in detail. Many rhabdomyosarcoma cells show normal expression of myogenic factors such as MyoD. In spite of this, the cells fail to arrest their growth and differentiate when exposed to signals, which should stimulate these events. Normally, RD cells were grown in medium containing 10% foetal bovine serum. When grown in medium containing only 1 % foetal bovine serum for several days, the RD cells spontaneously formed multinucleated syncitia (figure 22A). This process was, however, not very efficient and after 4 days in low-serum medium, only 5-10% of cells had undergone fusion.

Before studying the effect of overexpressing ADAM 12 in RD cells, we first looked at the endogenous localisation of ADAM 12, which was expressed in these rhabdomyosarcoma cells. To do this, RD cells were grown on glass coverslips and fixed with methanol/acetone in order to stain the cells by immunofluorescence using an anti-ADAM 12 polyclonal anti-serum. Interestingly we found that

ADAM 12 did not localise to the plasma membrane, as would be expected for a type I integral membrane protein (figure 22B).

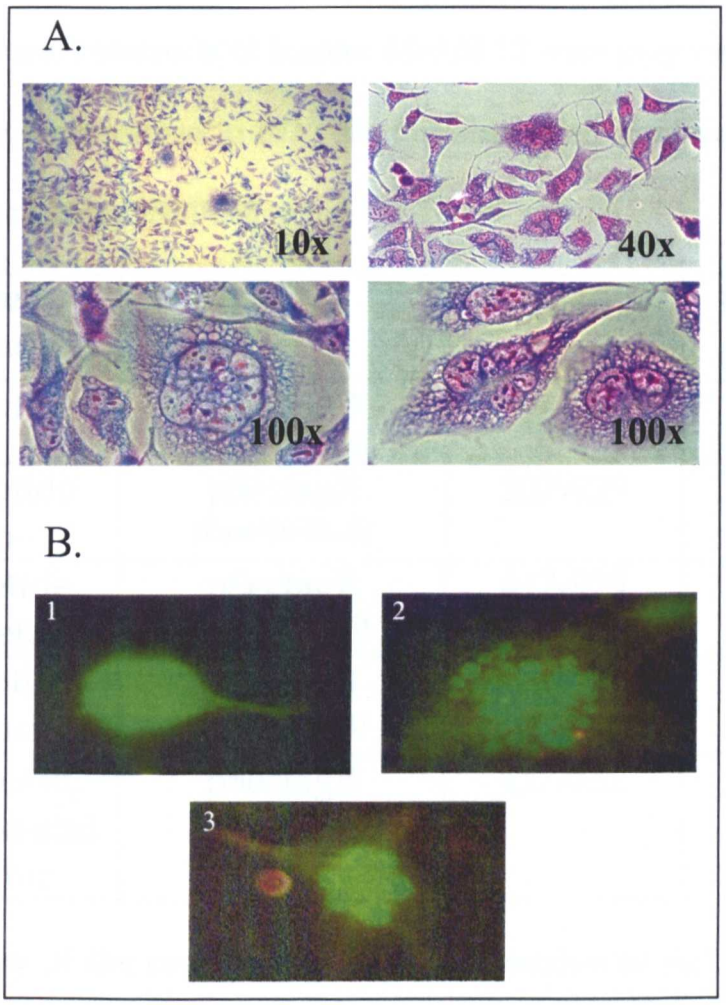

Figure 22: (A) Partially differentiated RD cells treated with the May-Grünwald and Giemsa stain to visualise the nuclei. (B) Immunofluorescence analysis using anti-ADAM 12 antibody of the endogenous localisation of ADAM 12 in single (1) or multinucleated (2 & 3) RD cells.

From the immunofluorescence studies, it appears that endogenous ADAM 12 localises to the cytoplasm and in a perinuclear compartment. If ADAM 12 is involved in the differentiation process of myoblasts and RD cells, its presence at the cell surface of cells may be regulated to ensure that fusion of cells does not occur randomly, but rather in response to a specific stimulus.

1.2 Immunofluorescence studies of ADAM 12

To further investigate the putative role of ADAM 12 in muscle cell development, several constructs of human ADAM 12 were prepared (summarised in Table 4) for overexpression in RD cells but also in an unrelated cell line, the human embryonic kidney 293T cell line.

	Truncated domains	Vector	Residues	Detectable expression
1	Signal peptide	pSectagB (EcoRV/XhoI)	29-929	Yes
2	Signal peptide and pro-	pSectagB (BamHI/XhoI)	207-929	Yes
3	Signal peptide, pro- and metallo-	pSectagB (BamHI/XhoI)	417-929	Yes
4	Signal peptide and cytoplasmic	pSectagB (EcoRV/XhoI)	29-732	Yes
5	Signal peptide, pro-, metallo-and cytoplasmic	pSectagB (BamHI/XhoI)	417-732	Yes

Table 4: Summary of the constructs used for expression of ADAM-12 in 293T and RD cells.

The pSecTagB (Invitrogen) vector is used for expression of genes under the control of the human cytomegalovirus (CMV) immediate-early promoter, and includes an Igκ-chain leader sequence, for efficient secretion of the recombinant protein and a C-terminal *myc*-epitope for detection with an anti-*myc* antibody.

Initial transfection results clearly showed that overexpression of either the full-length ADAM 12 (construct 1) or the expected active form of ADAM 12, missing the prodomain (construct 2), had no significant effect on the fusion ability of the transfected RD cells. No significant increase in the extent of syncitia formation could be observed.

All five constructs cloned into pSecTagB were used for cellular localisation studies of ADAM 12. Progressive deletion of the extracellular domains and the cytoplasmic domain was carried out in order to study the effect of each domain on cell localisation. These constructs were transfected into the muscle cell line, RD cells, but also into 293T cells (Figure 23), to see whether the cell localisation was tissue-dependent.

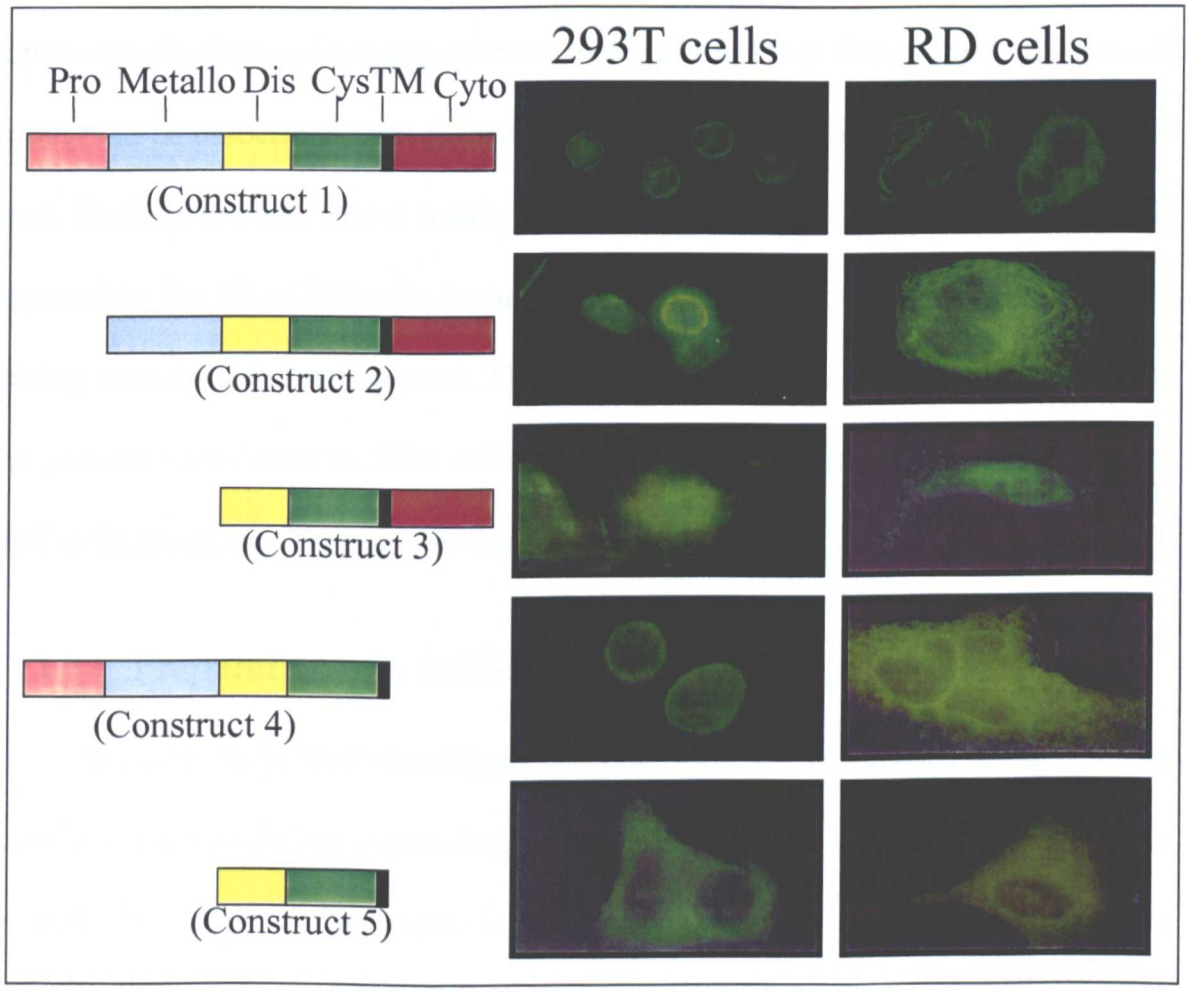

Figure 23: Immunofluorescence analysis of ADAM 12 localisation in 293T cells and RD cells. The anti-*myc* antibody was used for detection of the recombinant protein.

The results presented in figure 23 clearly demonstrate that the localisation of ADAM 12 differs from one cell line to another. In 293T cells, ADAM 12 visibly localises to the plasma membrane. Only in the absence of the metalloprotease domain, is the protein found in the endoplasmic reticulum (ER) and/or Golgi

apparatus. The cellular localisation was determined using known markers of the different subcellular compartments (a monoclonal anti-Golgi apparatus antibody (Ab-1), Oncogene; a monoclonal anti-nuclear pore antibody, Calbiochem; a monoclonal anti-rough endoplasmic reticulum antibody, Dako). This may result from a misfolding of the protein due to the missing domains, which leads to the retention of the protein in the ER. In the RD cells, however, none of the ADAM 12 constructs display a plasma membrane localisation. From the staining, the protein appears to be associated with microtubules and a perinuclear compartment, which most likely is the ER. These results would suggest that ADAM 12 is specifically retained in the ER of RD cells, in order to regulate its presence at the cell surface during muscle cell development. The fact that ADAM 12 is correctly targeted to the plasma membrane in 293T cells, also suggests that the RD cells and not the 293T cells are able to regulate the compartmentalisation of ADAM 12.

1.3 Preparation of a stable cell line expressing ADAM 12

In order to further investigate the potential role of ADAM 12 as a fusion protein, a stable cell line expressing ADAM 12 under an inducible promoter was created. The aim of this work was to have a cellular system in which the expression of ADAM 12 could be simultaneously induced in all cells. The presence of ADAM 12 in all cells, as opposed to only in singly transfected cells, may be required for its role in myoblast development. For this purpose the full-length ADAM 12 protein (residues 1-929) was cloned into pcDNA4/TO/Myc-HisC (EcoRI/XhoI sites; Invitrogen) for use with the T-Rex™ System (Invitrogen). This system is a tetracycline-regulated mammalian expression system that uses regulatory elements from the *E. coli* Tn*10*-encoded tetracycline (tet) resistance

operon. The system is based on the binding of tetracycline to the Tet repressor, which leads to a derepression of the promoter controlling the expression of the gene of interest. In addition to the pcDNA4/TO/Myc-HisC vector that contains the gene of interest, a second plasmid is needed, the pcDNA6/TR© regulatory vector, which expresses high levels of the TetR gene that encodes the tetracycline repressor.

Thus, to obtain the stable RD cell line expressing ADAM 12, a first round of selection of RD clones containing the pcDNA6/TR© was carried out, followed by a second round of selection for stable transformants containing both the pcDNA6/TR© vector and the pcDNA4/TO/Myc-HisC in which the ADAM 12 gene was cloned.

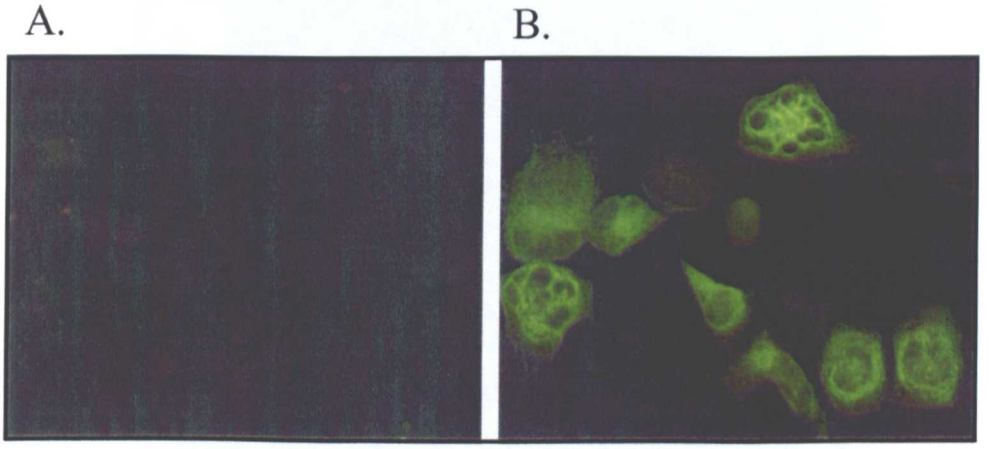

Figure 24: Indirect immunofluorescence analysis using an anti-*myc* antibody of stably transfected RD clones expressing full-length ADAM 12, in the absence (A) and presence (B) of the inducing agent, tetracycline.

The resulting clones were tested for their capacity to express ADAM 12 in response to the addition of 1μg/ml of tetracycline to the cell culture supernatant. A number of clones were obtained that efficiently expressed ADAM 12 in response to tetracycline (figure 24). However, as in the case of transient

transfections, ADAM 12 was restricted to a perinuclear localisation and the extent of syncitia formation in uninduced and induced cells were very similar.

Interestingly, however, this work did lead to the isolation of a RD clone (RD/TR 2.4 ℜ) that displayed a higher capacity to form syncitia.

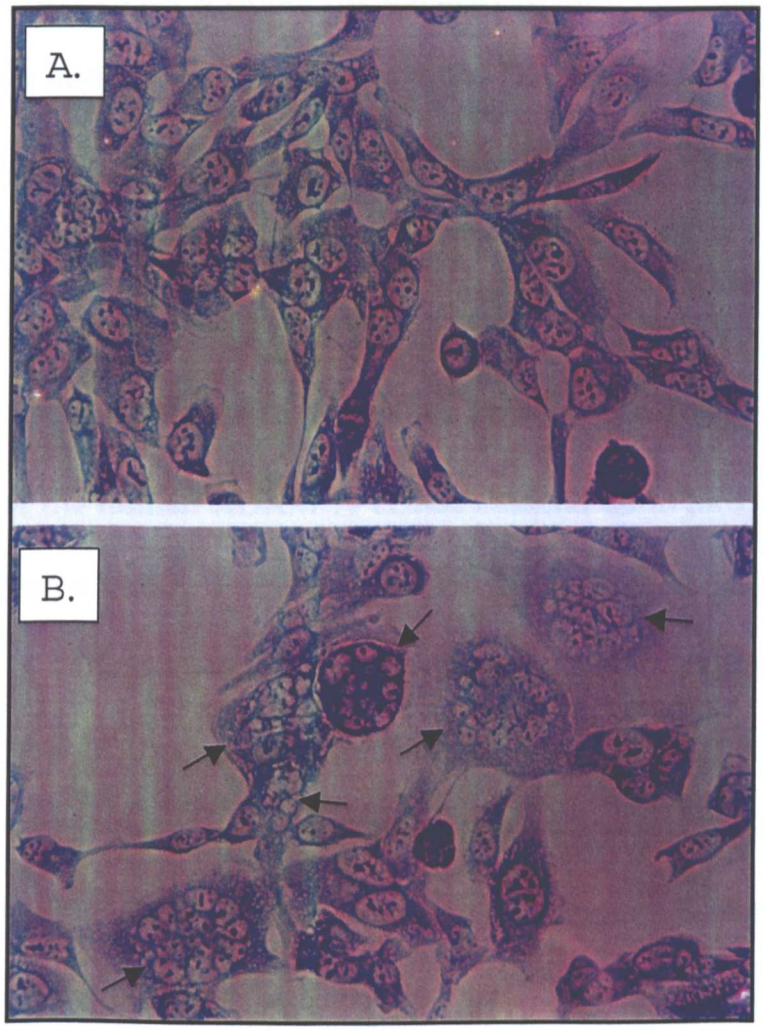

Figure 25: Giemsa and May-Grünwald staining of stably transfected RD clones expressing ADAM 12. (A) Clone RD/TR 2.4 ∅, which resembles the parental RD cell line, and (B) clone RD/TR 2.4 ℜ that forms syncitia (arrows) more efficiently.

Compared to the original RD cell line or to other stable transfectants (such as RD/TR 2.4 ∅) that form very few syncitia (less than 1% of cells) in serum-

containing medium, this new isolated RD clone could form up to 20-25% syncitia (figure 25). Both the RD/TR 2.4 ℜ and the RD/TR 2.4 ∅ clones were grown in the same conditions and displayed a strong expression of ADAM 12 in response to tetracycline. Expression of ADAM 12 was clearly not responsible for syncitia formation in the RD/TR 2.4 ℜ clone. This capacity to form multinucleated syncitia was spontaneous and not in response to starvation (incubating the cells in 1% foetal bovine serum; (Vachon et al., 1996)) or to the addition of differentiation agents, such as actinomycin D (Marchal et al., 1997).

2. Expression pattern of ADAM 19

Cloning of the human ADAM 19 cDNA enabled us to look into the tissue distribution pattern of this new ADAM protein. The expression pattern of ADAM 19 was studied in various tissues using the oligonucleotides designed to amplify the original 300 bp EST sequence corresponding to ADAM 19 (figure 26). The templates were either human cDNA libraries (in the case of skeletal and heart muscle, and uterus) or cDNA synthesised from RNA extracted from the tissue of patients (in the case of heart muscle and spleen). Interestingly, the human ADAM 19 cDNA, like its mouse homologue is expressed at variable levels in different tissues. Notably, although ADAM 19, like ADAM 12 is expressed at a high level in skeletal muscle, its expression is non-detectable in heart tissue. Since cardiac muscle cells, unlike skeletal myoblasts, do not undergo differentiation to form multinucleated myotubes, a potential role for ADAM 19 in the fusion of myoblasts is in agreement with this expression pattern. In addition, it was of interest to observe that when using cDNA synthesised from RNA extracted from patients

suffering from heart diseases, a faint band corresponding to the 300 bp EST fragment could be amplified (figure 26).

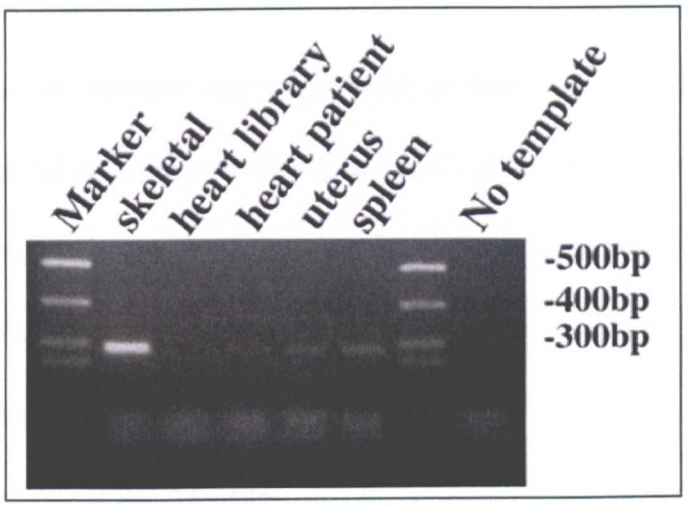

Figure 26: PCR analysis of the tissue distribution of ADAM 19. A 300 bp fragment was amplified corresponding to ADAM 19. The templates used for each PCR reaction are shown above each lane.

3. Discussion

In the functional approach to the study of ADAM 12, two main themes were addressed: the cellular localisation and targeting of ADAM 12 in RD cells, and the expression pattern of the newly characterised human ADAM 19. Analysis of the localisation of both the endogenous ADAM 12 and of recombinantly expressed ADAM 12 taught us that a specific, unknown, mechanism is involved in the retention of ADAM 12 in a perinuclear compartment in RD cells. This retention could not be abolished by the progressive deletion of multiple domains, suggesting that the signal for retention most probably comes from the RD cells and not from the primary sequence of the protein. However, all the constructs tested contained the transmembrane domain, so we cannot rule out the involvement of this domain in the retention mechanism. In 293T cells, ADAM 12

was properly targeted to the plasma membrane as would be expected for a type I integral membrane protein. The finding that ADAM 12 is specifically retained in a perinuclear compartment was described previously (Hougaard et al., 2000). In this work, Hougaard *et al.* studied the trafficking of full-length and several deletion mutants of ADAM 12 in COS-7, HeLa and CHO cells, by transient transfections. Similarly it was found that ADAM 12 accumulated in a perinuclear compartment, which was determined to be the *trans*-Golgi network. In contrast to our finding that ADAM 12 was correctly targeted to the plasma membrane in 293T cells, in this study ADAM 12 was retained in all three cell lines tested. In addition, although removal of various domains of ADAM 12 had no effect on the cellular localisation of ADAM 12 in RD cells, Hougaard *et al.* found that deletion of the cytoplasmic domain of ADAM 12 allowed for proper targeting of ADAM 12 to the plasma membrane. These results clearly show that the cell system used for the transient transfections is of crucial importance, since considerable differences can be observed. Hougaard *et al.* also found that the cytoplasmic domain alone was not sufficient to confer a perinuclear localisation to a chimeric Epidermal Growth Factor Receptor (EGFR; normally present at the cell surface) fused to the cytoplasmic domain of ADAM 12. A protein composed of the EGFR extracellular domains and the ADAM 12 transmembrane and cytoplasmic domains, nevertheless, was retained in a perinuclear compartment. Taken together, these results suggest that the transmembrane domain acts as a retention signal for ADAM 12, but not all types of cells are capable of recognising this signal. Restricted surface localisation has been described for other members of the ADAM family both *in vivo* and *in vitro*. This is the case for ADAM 2 (fertilin β) that only appears at the cell surface during the late stages of sperm maturation, although it

is already expressed in early spermatogenesis (Carroll et al., 1995). A reservoir of ADAM 12 in a perinuclear compartment would provide cells with a mechanism for rapid and tightly regulated release of ADAM 12 to the cell surface in response to specific stimuli. At the cell surface, ADAM 12 would then carry out its functions in proteolytic degradation and/or adhesion.

From this work, it was also very clear that ADAM 12 alone was not sufficient to provoke myoblast fusion, even when ADAM 12 was expressed at the cell surface of cells (as in 293T cells and in Sf9 insect cells). These results strongly imply that this developmental process involves multiple proteins, and that the overexpression of one of these is not sufficient to promote myoblast fusion.

The cloning of the human ADAM 19 cDNA enabled us to investigate the tissue distribution of a protein closely related to ADAM 12. Its high level of expression in skeletal muscle and its absence in cardiac muscle support a possible role of ADAM 19 in myoblast differentiation, in conjunction with ADAM 12. A similar expression pattern was also observed for mouse ADAM 19 (Yagami-Hiromasa et al., 1995), whereas more recent work by Wei *et al.* suggested that human ADAM 19 was expressed in a wide variety of tissues, including heart (Wei et al., 2001). This discrepancy may be due to the different techniques used to study the expression (Northern blotting versus the more sensitive RT-PCR). In addition to the PCR results, which clearly show that ADAM 19 is not present at a detectable level in heart, a trial to screen an amplified heart muscle cDNA library using a radioactively labelled EST fragment (corresponding to the end of the disintegrin domain and the beginning of the cysteine-rich domain) failed to isolate positive clones of ADAM 19. In our case, the results suggest that the expression of ADAM 19 is most likely tissue-dependent and developmentally regulated. The finding

that ADAM 19 is expressed at a low level in patients suffering from heart diseases, but not in normal heart tissue, strongly suggests that ADAM 19 may be involved in the development and progression of the disease or in the cellular response to this disease state. It is quite common to observe that genes that are normally switched off in a given tissue become expressed in the same diseased tissue. Sarcoplasmic reticulum genes are upregulated in mild cardiac hypertrophy but downregulated in severe cardiac hypertrophy induced by pressure overload (Arai et al., 1996). A 2.5-fold increase in HOXA5 (a member of the HOX gene family of transcription factors) mRNA expression was demonstrated by quantitative reverse transcriptase-polymerase chain reaction in primary pulmonary hypertension lung specimens when compared to normal lung tissue (Golpon et al., 2001). Differences in the pattern of HOX gene expression exist among foetal, adult, and diseased lung specimens and this altered pattern of gene expression may contribute to the development of pulmonary diseases. Riemer *et al.* have studied the effect of scrapie-infected brains on gene regulation. Here again, it was found by subtractive hybridisation, northern blotting and nucleotide sequencing that 19 genes are up-regulated in scrapie-infected brain tissue, relative to normal brain tissue (Riemer et al., 2000). Many tissues display considerable changes in their gene expression during disease, and the identification of the genes with altered expression patterns may provide further insight on the molecular level into the processes underlying these diseases.

Like ADAM 12, ADAM 19 has recently been shown to be an active metalloprotease (Wei et al., 2001), based on the sequence homology of the zinc-binding site to other metalloproteinases and the finding that ADAM 19 can form a complex with α2-macroglobulin. However, its biological substrate *in vivo* still

remains to be determined. Further work will be needed to shed light on the roles of ADAM 12 and ADAM 19 *in vivo* and to find out whether the heterodimerisation of ADAM 12 and ADAM 19 has a biological function in myoblast development, in a similar way to the fertilin complex in fertilisation. In addition, a better understanding of the processes leading to fusion of myoblasts will also require the identification of new molecules, and more particularly cell surface proteins, that may also be involved in this complex developmental process. We are currently collaborating with a laboratory at the Commissariat à l'Energie Atomique (CEA) to analyse the proteins present at the plasma membrane of the parental RD cell line and of the new syncitia-forming RD cell line (RD/TR 2.4 ℜ). The aim of this work is to isolate new cell surface molecules that are required for efficient fusion of myoblasts, and thus would be found in the new isolated cell line, but not in the original one. Isolated plasma membrane preparations from these two cell lines are currently being used for high-throughput mass spectrometry analysis. Although this type of large-scale analysis requires considerable optimisation of the quality of the plasma membrane preparations in order to avoid contamination, this work should lead to the identification of new, important integral membrane or membrane-associated proteins that are directly involved in the formation of multinucleated myotubes.

References:

Abe, E., Mocharla, H., Yamate, T., Taguchi, Y., and Manolagas, S. C. (1999). Meltrin-alpha, a fusion protein involved in multinucleated giant cell and osteoclast formation. Calcif Tissue Int *64*, 508-515.

Abe, Y., Odaka, M., Inagaki, F., Lax, I., Schlessinger, J., and Kohda, D. (1998). Disulfide bond structure of human epidermal growth factor receptor. J Biol Chem 273, 11150-11157.

Allison, S. L., Schalich, J., Stiasny, K., Mandl, C. W., and Heinz, F. X. (2001). Mutational evidence for an internal fusion peptide in flavivirus envelope protein E. J Virol *75*, 4268-4275.

Almeida, E. A., Huovila, A. P., Sutherland, A. E., Stephens, L. E., Calarco, P. G., Shaw, L. M., Mercurio, A. M., Sonnenberg, A., Primakoff, P., Myles, D. G., and et al. (1995). Mouse egg integrin alpha 6 beta 1 functions as a sperm receptor. Cell *81*, 1095-1104.

Arai, M., Suzuki, T., and Nagai, R. (1996). Sarcoplasmic reticulum genes are upregulated in mild cardiac hypertrophy but downregulated in severe cardiac hypertrophy induced by pressure overload. J Mol Cell Cardiol *28*, 1583-1590.

Asakura, M., Kitakaze, M., Takashima, S., Liao, Y., Ishikura, F., Yoshinaka, T., Ohmoto, H., Node, K., Yoshino, K., Ishiguro, H., Asanuma, H., Sanada, S., Matsumura, Y., Takeda, H., Beppu, S., Tada, M., Hori, M., and Higashiyama, S. (2002). Cardiac hypertrophy is inhibited by antagonism of ADAM12 processing of HB-EGF: Metalloproteinase inhibitors as a new therapy. Nat Med *8*, 35-40.

Bijlenga, P., Liu, J. H., Espinos, E., Haenggeli, C. A., Fischer-Lougheed, J., Bader, C. R., and Bernheim, L. (2000). T-type alpha 1H Ca2+ channels are involved in Ca2+ signaling during terminal differentiation (fusion) of human myoblasts. Proc Natl Acad Sci U S A *97*, 7627-7632.

Black, B. L., and Olson, E. N. (1998). Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. Annu Rev Cell Dev Biol *14*, 167-196.

Black, R. A., Rauch, C. T., Kozlosky, C. J., Peschon, J. J., Slack, J. L., Wolfson, M. F., Castner, B. J., Stocking, K. L., Reddy, P., Srinivasan, S., Nelson, N., Boiani, N., Schooley, K. A., Gerhart, M., Davis, R., Fitzner, J. N., Johnson, R. S., Paxton, R. J., March, C. J., and Cerretti, D. P. (1997). A metalloproteinase disintegrin that releases tumour-necrosis factor- alpha from cells. Nature *385*, 729-733.

Black, R. A., and White, J. M. (1998). ADAMs: focus on the protease domain. Curr Opin Cell Biol *10*, 654-659.

Blau, H. M., Dhawan, J., and Pavlath, G. K. (1993). Myoblasts in pattern formation and gene therapy. Trends Genet *9*, 269-274.

Blobel, C. P., Myles, D. G., Primakoff, P., and White, J. M. (1990). Proteolytic processing of a protein involved in sperm-egg fusion correlates with acquisition of fertilization competence. J Cell Biol *111*, 69-78.

Blobel, C. P., Wolfsberg, T. G., Turck, C. W., Myles, D. G., Primakoff, P., and White, J. M. (1992). A potential fusion peptide and an integrin ligand domain in a protein active in sperm-egg fusion. Nature *356*, 248-252.

Bour, B. A., Chakravarti, M., West, J. M., and Abmayr, S. M. (2000). Drosophila SNS, a member of the immunoglobulin superfamily that is essential for myoblast fusion. Genes Dev *14*, 1498-1511.

Brickman, Y. G., Ford, M. D., Small, D. H., Bartlett, P. F., and Nurcombe, V. (1995). Heparan sulfates mediate the binding of basic fibroblast growth factor to a specific receptor on neural precursor cells. J Biol Chem *270*, 24941-24948.

Cal, S., Freije, J. M., Lopez, J. M., Takada, Y., and Lopez-Otin, C. (2000). ADAM 23/MDC3, a human disintegrin that promotes cell adhesion via interaction with the alphavbeta3 integrin through an RGD-independent mechanism. Mol Biol Cell *11*, 1457-1469.

Calvete, J. J., Schrader, M., Raida, M., McLane, M. A., Romero, A., and Niewiarowski, S. (1997). The disulphide bond pattern of bitistatin, a disintegrin isolated from the venom of the viper Bitis arietans. FEBS Lett *416*, 197-202.

Calvete, J. J., Moreno-Murciano, M. P., Sanz, L., Jurgens, M., Schrader, M., Raida, M., Benjamin, D. C., and Fox, J. W. (2000). The disulfide bond pattern of catrocollastatin C, a disintegrin-like/cysteine-rich protein isolated from Crotalus atrox venom. Protein Sci *9*, 1365-1373.

Carey, D. J. (1997). Syndecans: multifunctional cell-surface co-receptors. Biochem J *327*, 1-16.

Carroll, D. J., Dikegoros, E., Koppel, D. E., and Cowan, A. E. (1995). Surface expression of the pre-beta subunit of fertilin is regulated at a post-translational level in guinea pig spermatids. Dev Biol *168*, 429-437.

Chen, E. H., and Olson, E. N. (2001). Antisocial, an intracellular adaptor protein, is required for myoblast fusion in Drosophila. Dev Cell *1*, 705-715.

Chen, H., and Sampson, N. S. (1999). Mediation of sperm-egg fusion: evidence that mouse egg alpha6beta1 integrin is the receptor for sperm fertilinbeta. Chem Biol *6*, 1-10.

Cho, C., Bunch, D. O., Faure, J. E., Goulding, E. H., Eddy, E. M., Primakoff, P., and Myles, D. G. (1998). Fertilization defects in sperm from mice lacking fertilin beta. Science *281*, 1857-1859.

Crawley, S., Farrell, E. M., Wang, W., Gu, M., Huang, H. Y., Huynh, V., Hodges, B. L., Cooper, D. N., and Kaufman, S. J. (1997). The alpha7beta1 integrin mediates adhesion and migration of skeletal myoblasts on laminin. Exp Cell Res *235*, 274-286.

Doberstein, S. K., Fetter, R. D., Mehta, A. Y., and Goodman, C. S. (1997). Genetic analysis of myoblast fusion: blown fuse is required for progression beyond the prefusion complex. J Cell Biol *136*, 1249-1261.

Duan, H., Skeath, J. B., and Nguyen, H. T. (2001). Drosophila Lame duck, a novel member of the Gli superfamily, acts as a key regulator of myogenesis by controlling fusion-competent myoblast development. Development *128*, 4489-4500.

Dworak, H. A., Charles, M. A., Pellerano, L. B., and Sink, H. (2001). Characterization of Drosophila hibris, a gene related to human nephrin. Development *128*, 4265-4276.

Erickson, M. R., Galletta, B. J., and Abmayr, S. M. (1997). Drosophila myoblast city encodes a conserved protein that is essential for myoblast fusion, dorsal closure, and cytoskeletal organization. J Cell Biol *138*, 589-603.

Eto, K., Puzon-McLaughlin, W., Sheppard, D., Sehara-Fujisawa, A., Zhang, X. P., and Takada, Y. (2000). RGD-independent binding of integrin alpha9beta1 to the ADAM-12 and -15 disintegrin domains mediates cell-cell interaction. J Biol Chem *275*, 34922-34930.

Evans, J. P. (1999). Sperm disintegrins, egg integrins, and other cell adhesion molecules of mammalian gamete plasma membrane interactions. Front Biosci *4*, D114-131.

Evans, J. P. (2001). Fertilin beta and other ADAMs as integrin ligands: insights into cell adhesion and fertilization. Bioessays *23*, 628-639.

Ferris, P. J., Woessner, J. P., and Goodenough, U. W. (1996). A sex recognition glycoprotein is encoded by the plus mating-type gene fus1 of Chlamydomonas reinhardtii. Mol Biol Cell *7*, 1235-1248.

Fischer-Lougheed, J., Liu, J. H., Espinos, E., Mordasini, D., Bader, C. R., Belin, D., and Bernheim, L. (2001). Human myoblast fusion requires expression of functional inward rectifier Kir2.1 channels. J Cell Biol *153*, 677-686.

Fuentealba, L., Carey, D. J., and Brandan, E. (1999). Antisense inhibition of syndecan-3 expression during skeletal muscle differentiation accelerates myogenesis through a basic fibroblast growth factor-dependent mechanism. J Biol Chem *274*, 37876-37884.

Gilpin, B. J., Loechel, F., Mattei, M. G., Engvall, E., Albrechtsen, R., and Wewer, U. M. (1998). A novel, secreted form of human ADAM 12 (meltrin alpha) provokes myogenesis in vivo. J Biol Chem *273*, 157-166.

Goichberg, P., and Geiger, B. (1998). Direct involvement of N-cadherin-mediated signaling in muscle differentiation. Mol Biol Cell *9*, 3119-3131.

Goichberg, P., Shtutman, M., Ben-Ze'ev, A., and Geiger, B. (2001). Recruitment of beta-catenin to cadherin-mediated intercellular adhesions is involved in myogenic induction. J Cell Sci *114*, 1309-1319.

Golpon, H. A., Geraci, M. W., Moore, M. D., Miller, H. L., Miller, G. J., Tuder, R. M., and Voelkel, N. F. (2001). HOX genes in human lung: altered expression in primary pulmonary hypertension and emphysema. Am J Pathol *158*, 955-966.

Gullberg, D., Sjoberg, G., Velling, T., and Sejersen, T. (1995). Analysis of fibronectin and vitronectin receptors on human fetal skeletal muscle cells upon differentiation. Exp Cell Res *220*, 112-123.

Harris, H. A., Murrills, R. J., and Komm, B. S. (1997). Expression of meltrin-alpha mRNA is not restricted to fusagenic cells. J Cell Biochem *67*, 136-142.

Hernandez, L. D., Hoffman, L. R., Wolfsberg, T. G., and White, J. M. (1996). Virus-cell and cell-cell fusion. Annu Rev Cell Dev Biol *12*, 627-661.

Higgins, D. and Cregg, JM (1998). *Pichia* Protocols, Vol 103 (Totowa, New Jersey, Humana Press).

Hite, L. A., Jia, L. G., Bjarnason, J. B., and Fox, J. W. (1994). cDNA sequences for four snake venom metalloproteinases: structure, classification, and their relationship to mammalian reproductive proteins. Arch Biochem Biophys *308*, 182-191.

Hooft van Huijsduijnen, R. (1998). ADAM 20 and 21; two novel human testis-specific membrane metalloproteases with similarity to fertilin-alpha. Gene *206*, 273-282.

Hougaard, S., Loechel, F., Xu, X., Tajima, R., Albrechtsen, R., and Wewer, U. M. (2000). Trafficking of human ADAM 12-L: retention in the trans-Golgi network. Biochem Biophys Res Commun *275*, 261-267.

Hynes, R. O. (1987). Integrins: a family of cell surface receptors. Cell *48*, 549-554.

Iba, K., Albrechtsen, R., Gilpin, B. J., Loechel, F., and Wewer, U. M. (1999). Cysteine-rich domain of human ADAM 12 (meltrin alpha) supports tumor cell adhesion. Am J Pathol *154*, 1489-1501.

Iba, K., Albrechtsen, R., Gilpin, B., Frohlich, C., Loechel, F., Zolkiewska, A., Ishiguro, K., Kojima, T., Liu, W., Langford, J. K., Sanderson, R. D., Brakebusch, C., Fassler, R., and Wewer, U. M. (2000). The cysteine-rich domain of human ADAM 12 supports cell adhesion through syndecans and triggers signaling events that lead to beta1 integrin-dependent cell spreading. J Cell Biol *149*, 1143-1156.

Inoue, D., Reid, M., Lum, L., Kratzschmar, J., Weskamp, G., Myung, Y. M., Baron, R., and Blobel, C. P. (1998). Cloning and initial characterization of mouse meltrin beta and analysis of the expression of four metalloprotease-disintegrins in bone cells. J Biol Chem *273*, 4180-4187.

Jia, L. G., Shimokawa, K., Bjarnason, J. B., and Fox, J. W. (1996). Snake venom metalloproteinases: structure, function and relationship to the ADAMs family of proteins. Toxicon *34*, 1269-1276.

Johnson, W. C., Jr. (1990). Protein secondary structure and circular dichroism: a practical guide. Proteins *7*, 205-214.

Kang, J. S., Mulieri, P. J., Hu, Y., Taliana, L., and Krauss, R. S. (2002). BOC, an Ig superfamily member, associates with CDO to positively regulate myogenic differentiation. Embo J *21*, 114-124.

Kang, Q., Cao, Y., and Zolkiewska, A. (2000). Metalloprotease-disintegrin ADAM 12 binds to the SH3 domain of Src and activates Src tyrosine kinase in C2C12 cells. Biochem J *352 Pt 3*, 883-892.

Kang, Q., Cao, Y., and Zolkiewska, A. (2001). Direct interaction between the cytoplasmic tail of ADAM 12 and the Src homology 3 domain of p85alpha activates phosphatidylinositol 3-kinase in C2C12 cells. J Biol Chem *276*, 24466-24472.

Kim, D., Chi, S., Lee, K. H., Rhee, S., Kwon, Y. K., Chung, C. H., Kwon, H., and Kang, M. S. (1999). Neuregulin stimulates myogenic differentiation in an autocrine manner. J Biol Chem *274*, 15395-15400.

Knudsen, K. A., and Horwitz, A. F. (1977). Tandem events in myoblast fusion. Dev Biol *58*, 328-338.

Knudsen, K. A., and Horwitz, A. F. (1978). Toward a mechanism of myoblast fusion. Prog Clin Biol Res *23*, 563-568.

Knudsen, K. A. (1992). Membrane fusion, Marcel Dekker, Inc., New York.

Kratzschmar, J., Lum, L., and Blobel, C. P. (1996). Metargidin, a membrane-anchored metalloprotease-disintegrin protein with an RGD integrin binding sequence. J Biol Chem *271*, 4593-4596.

Kuno, K., Iizasa, H., Ohno, S., and Matsushima, K. (1997). The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene encoding an ADAM family protein with TSP motifs. Genomics *46*, 466-471.

Lammich, S., Kojro, E., Postina, R., Gilbert, S., Pfeiffer, R., Jasionowski, M., Haass, C., and Fahrenholz, F. (1999). Constitutive and regulated alpha-secretase cleavage of Alzheimer's amyloid precursor protein by a disintegrin metalloprotease. Proc Natl Acad Sci U S A *96*, 3922-3927.

Lewis, M. P., Tippett, H. L., Sinanan, A. C., Morgan, M. J., and Hunt, N. P. (2000). Gelatinase-B (matrix metalloproteinase-9; MMP-9) secretion is involved in the migratory phase of human and murine muscle cell cultures. J Muscle Res Cell Motil *21*, 223-233.

Li, J., Mayne, R., and Wu, C. (1999). A novel muscle-specific beta 1 integrin binding protein (MIBP) that modulates myogenic differentiation. J Cell Biol *147*, 1391-1398.

Loechel, F., Gilpin, B. J., Engvall, E., Albrechtsen, R., and Wewer, U. M. (1998). Human ADAM 12 (meltrin alpha) is an active metalloprotease. J Biol Chem *273*, 16993-16997.

Loechel, F., Overgaard, M. T., Oxvig, C., Albrechtsen, R., and Wewer, U. M. (1999). Regulation of human ADAM 12 protease by the prodomain. Evidence for a functional cysteine switch. J Biol Chem *274*, 13427-13433.

Loechel, F., Fox, J. W., Murphy, G., Albrechtsen, R., and Wewer, U. M. (2000). ADAM 12-S cleaves IGFBP-3 and IGFBP-5 and is inhibited by TIMP-3. Biochem Biophys Res Commun *278*, 511-515.

Loechel, F., and Wewer, U. M. (2001). Activation of ADAM 12 protease by copper. FEBS Lett *506*, 65-68.

Lunn, C. A., Fan, X., Dalie, B., Miller, K., Zavodny, P. J., Narula, S. K., and Lundell, D. (1997). Purification of ADAM 10 from bovine spleen as a TNFalpha convertase. FEBS Lett *400*, 333-335.

Marchal, J. A., Prados, J., Melguizo, C., Fernandez, J. E., Velez, C., Alvarez, L., and Aranega, A. (1997). Actinomycin D treatment leads to differentiation and inhibits proliferation in rhabdomyosarcoma cells. J Lab Clin Med *130*, 42-50.

Martin, I., and Ruysschaert, J. M. (1997). Comparison of lipid vesicle fusion induced by the putative fusion peptide of fertilin (a protein active in sperm-egg fusion) and the NH2- terminal domain of the HIV2 gp41. FEBS Lett *405*, 351-355.

Martin, I., Epand, R. M., and Ruysschaert, J. M. (1998). Structural properties of the putative fusion peptide of fertilin, a protein active in sperm-egg fusion, upon interaction with the lipid bilayer. Biochemistry *37*, 17030-17039.

Maskos, K., Fernandez-Catalan, C., Huber, R., Bourenkov, G. P., Bartunik, H., Ellestad, G. A., Reddy, P., Wolfson, M. F., Rauch, C. T., Castner, B. J., Davis, R., Clarke, H. R., Petersen, M., Fitzner, J. N., Cerretti, D. P., March, C. J., Paxton, R. J., Black, R. A., and Bode, W. (1998). Crystal structure of the catalytic domain of human tumor necrosis factor-alpha-converting enzyme. Proc Natl Acad Sci U S A *95*, 3408-3412.

Matsumura, M., Signor, G., and Matthews, B. W. (1989). Substantial increase of protein stability by multiple disulphide bonds. Nature *342*, 291-293.

McKinsey, T. A., Zhang, C. L., and Olson, E. N. (2001). Control of muscle development by dueling HATs and HDACs. Curr Opin Genet Dev *11*, 497-504.

Megeney, L. A., and Rudnicki, M. A. (1995). Determination versus differentiation and the MyoD family of transcription factors. Biochem Cell Biol *73*, 723-732.

Menon, S. D., and Chia, W. (2001). Drosophila rolling pebbles: a multidomain protein required for myoblast fusion that recruits D-Titin in response to the myoblast attractant Dumbfounded. Dev Cell *1*, 691-703.

Myles, D. G., Kimmel, L. H., Blobel, C. P., White, J. M., and Primakoff, P. (1994). Identification of a binding site in the disintegrin domain of fertilin required for sperm-egg fusion. Proc Natl Acad Sci U S A *91*, 4195-4198.

Nath, D., Slocombe, P. M., Stephens, P. E., Warn, A., Hutchinson, G. R., Yamada, K. M., Docherty, A. J., and Murphy, G. (1999). Interaction of metargidin (ADAM-15) with alphavbeta3 and alpha5beta1 integrins on different haemopoietic cells. J Cell Sci *112*, 579-587.

Nath, D., Slocombe, P. M., Webster, A., Stephens, P. E., Docherty, A. J., and Murphy, G. (2000). Meltrin gamma(ADAM-9) mediates cellular adhesion through alpha(6)beta(1)integrin, leading to a marked induction of fibroblast cell motility. J Cell Sci *113*, 2319-2328.

Paine, M. J., Desmond, H. P., Theakston, R. D., and Crampton, J. M. (1992). Purification, cloning, and molecular characterization of a high molecular weight hemorrhagic metalloprotease, jararhagin, from Bothrops jararaca venom. Insights into the disintegrin gene family. J Biol Chem *267*, 22869-22876.

Pan, D., and Rubin, G. M. (1997). Kuzbanian controls proteolytic processing of Notch and mediates lateral inhibition during Drosophila and vertebrate neurogenesis. Cell *90*, 271-280.

Paululat, A., Burchard, S., and Renkawitz-Pohl, R. (1995). Fusion from myoblasts to myotubes is dependent on the rolling stone gene (rost) of Drosophila. Development *121*, 2611-2620.

Paululat, A., Goubeaud, A., Damm, C., Knirr, S., Burchard, S., and Renkawitz-Pohl, R. (1997). The mesodermal expression of rolling stone (rost) is essential for myoblast fusion in Drosophila and encodes a potential transmembrane protein. J Cell Biol *138*, 337-348.

Paululat, A., Holz, A., and Renkawitz-Pohl, R. (1999). Essential genes for myoblast fusion in Drosophila embryogenesis. Mech Dev *83*, 17-26.

Pirskanen, A., Kiefer, J. C., and Hauschka, S. D. (2000). IGFs, insulin, Shh, bFGF, and TGF-beta1 interact synergistically to promote somite myogenesis in vitro. Dev Biol *224*, 189-203.

Primakoff, P., Hyatt, H., and Tredick-Kline, J. (1987). Identification and purification of a sperm surface protein with a potential role in sperm-egg membrane fusion. J Cell Biol *104*, 141-149.

Primakoff, P., and Myles, D. G. (2000). The ADAM gene family: surface proteins with adhesion and protease activity. Trends Genet *16*, 83-87.

Rau, A., Buttgereit, D., Holz, A., Fetter, R., Doberstein, S. K., Paululat, A., Staudt, N., Skeath, J., Michelson, A. M., and Renkawitz-Pohl, R. (2001). rolling pebbles (rols) is required in Drosophila muscle precursors for recruitment of myoblasts for fusion. Development *128*, 5061-5073.

Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C., and Harrison, S. C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2 A resolution. Nature *375*, 291-298.

Riemer, C., Queck, I., Simon, D., Kurth, R., and Baier, M. (2000). Identification of upregulated genes in scrapie-infected brain tissue. J Virol *74*, 10245-10248.

Rohwedel, J., Guan, K., Zuschratter, W., Jin, S., Ahnert-Hilger, G., Furst, D., Fassler, R., and Wobus, A. M. (1998). Loss of beta1 integrin function results in a retardation of myogenic, but an acceleration of neuronal, differentiation of embryonic stem cells in vitro. Dev Biol *201*, 167-184.

Rosen, G. D., Sanes, J. R., LaChance, R., Cunningham, J. M., Roman, J., and Dean, D. C. (1992). Roles for the integrin VLA-4 and its counter receptor VCAM-1 in myogenesis. Cell *69*, 1107-1119.

Ruiz-Gomez, M., Coutts, N., Price, A., Taylor, M. V., and Bate, M. (2000). Drosophila dumbfounded: a myoblast attractant essential for fusion. Cell *102*, 189-198.

Rushton, E., Drysdale, R., Abmayr, S. M., Michelson, A. M., and Bate, M. (1995). Mutations in a novel gene, myoblast city, provide evidence in support of the founder cell hypothesis for Drosophila muscle development. Development *121*, 1979-1988.

Saginario, C., Qian, H. Y., and Vignery, A. (1995). Identification of an inducible surface molecule specific to fusing macrophages. Proc Natl Acad Sci U S A *92*, 12210-12214.

Saphire, A. C., Bobardt, M. D., Zhang, Z., David, G., and Gallay, P. A. (2001). Syndecans serve as attachment receptors for human immunodeficiency virus type 1 on macrophages. J Virol *75*, 9187-9200.

Spencer, J. A., Eliazer, S., Ilaria, R. L., Jr., Richardson, J. A., and Olson, E. N. (2000). Regulation of microtubule dynamics and myogenic differentiation by MURF, a striated muscle RING-finger protein. J Cell Biol *150*, 771-784.

Stronach, B. E., Renfranz, P. J., Lilly, B., and Beckerle, M. C. (1999). Muscle LIM proteins are associated with muscle sarcomeres and require dMEF2 for their expression during Drosophila myogenesis. Mol Biol Cell *10*, 2329-2342.

Suzuki, A., Kadota, N., Hara, T., Nakagami, Y., Izumi, T., Takenawa, T., Sabe, H., and Endo, T. (2000). Meltrin alpha cytoplasmic domain interacts with SH3 domains of Src and Grb2 and is phosphorylated by v-Src. Oncogene *19*, 5842-5850.

Tachibana, I., and Hemler, M. E. (1999). Role of transmembrane 4 superfamily (TM4SF) proteins CD9 and CD81 in muscle cell fusion and myotube maintenance. J Cell Biol *146*, 893-904.

Trueheart, J., and Fink, G. R. (1989). The yeast cell fusion protein FUS1 is O-glycosylated and spans the plasma membrane. Proc Natl Acad Sci U S A *86*, 9916-9920.

Vachon, P. H., Loechel, F., Xu, H., Wewer, U. M., and Engvall, E. (1996). Merosin and laminin in myogenesis; specific requirement for merosin in myotube stability and survival. J Cell Biol *134*, 1483-1497.

Van Wart, H. E., and Birkedal-Hansen, H. (1990). The cysteine switch: a principle of regulation of metalloproteinase activity with potential applicability to the entire matrix metalloproteinase gene family. Proc Natl Acad Sci U S A *87*, 5578-5582.

Wei, P., Zhao, Y. G., Zhuang, L., Ruben, S., and Sang, Q. X. (2001). Expression and enzymatic activity of human disintegrin and metalloproteinase ADAM19/meltrin beta. Biochem Biophys Res Commun *280*, 744-755.

Wen, C., Metzstein, M. M., and Greenwald, I. (1997). SUP-17, a Caenorhabditis elegans ADAM protein related to Drosophila KUZBANIAN, and its role in LIN-12/NOTCH signalling. Development *124*, 4759-4767.

White, J. M., and Blobel, C. P. (1989). Cell-to-cell fusion. Curr Opin Cell Biol *1*, 934-939.

White, J. M. (1992). Membrane fusion. Science *258*, 917-924.

Wolfsberg, T. G., Primakoff, P., Myles, D. G., and White, J. M. (1995). ADAM, a novel family of membrane proteins containing A Disintegrin And Metalloprotease domain: multipotential functions in cell-cell and cell- matrix interactions. J Cell Biol *131*, 275-278.

Wolfsberg, T. G., and White, J. M. (1996). ADAMs in fertilization and development. Dev Biol *180*, 389-401.

Wong, J. Y., Kuhl, T. L., Israelachvili, J. N., Mullah, N., and Zalipsky, S. (1997). Direct measurement of a tethered ligand-receptor interaction potential. Science *275*, 820-822.

Woods, A., and Couchman, J. R. (1998). Syndecans: synergistic activators of cell adhesion. Trends Cell Biol *8*, 189-192.

Yagami-Hiromasa, T., Sato, T., Kurisaki, T., Kamijo, K., Nabeshima, Y., and Fujisawa-Sehara, A. (1995). A metalloprotease-disintegrin participating in myoblast fusion. Nature *377*, 652-656.

Yuan, R., Primakoff, P., and Myles, D. G. (1997). A role for the disintegrin domain of cyritestin, a sperm surface protein belonging to the ADAM family, in mouse sperm-egg plasma membrane adhesion and fusion. J Cell Biol *137*, 105-112.

Zhang, Q., Skepper, J. N., Yang, F., Davies, J. D., Hegyi, L., Roberts, R. G., Weissberg, P. L., Ellis, J. A., and Shanahan, C. M. (2001). Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues. J Cell Sci *114*, 4485-4498.

Zhang, X. P., Kamata, T., Yokoyama, K., Puzon-McLaughlin, W., and Takada, Y. (1998). Specific interaction of the recombinant disintegrin-like domain of MDC- 15 (metargidin, ADAM-15) with integrin alphavbeta3. J Biol Chem *273*, 7345-7350.

Zhou, M., Graham, R., Russell, G., and Croucher, P. I. (2001). MDC-9 (ADAM-9/Meltrin gamma) functions as an adhesion molecule by binding the alpha(v)beta(5) integrin. Biochem Biophys Res Commun *280*, 574-580.

Part II

Role of the Ebola Virus Matrix Protein, VP40,

In Virus Assembly and Release

Chapter 4: Introduction

1. Ebola virus

1.1 The pathogenesis of Ebola haemorrhagic fever

1.1.1 A deadly pathogen

Ebola and Marburg viruses are classified as members of the family Filoviridae (Kiley et al., 1982). The first member of the Filoviridae to be described was Marburg virus, which appeared in Marburg, Germany, in 1967 (Martini, 1971), while Ebola virus was first described in Central Africa in 1976. A first outbreak occurred in July 1976 in Sudan and two months later in the Democratic Republic of the Congo, causing many deaths (280 deaths out of 318 cases) (Johnson et al., 1977; Samaranayake et al., 1996). Several other outbreaks then occurred in the following years (1977-1979) in this same area. In 1989, Ebola appeared in monkeys imported from the Philippines into a Reston, Virginia (USA), primate facility. Epidemics in cynomolgus monkeys (*Macaca fascicularis*) occurred in this facility and others throughout 1992 (Jahrling et al., 1990) and recurred in 1996. This virus strain (Reston), however, appeared to have a reduced pathogenicity compared to the Zaire Ebola subtype that caused the numerous deaths in Central Africa. At present, Ebola virus is subdivided into four different subtypes: Zaire, Sudan, Ivory Coast and Reston (Peters, 1996). The subtypes Ebola Sudan and Reston are less virulent and cause a self-limiting infection in monkeys (Fisher-Hoch et al., 1992). Ebola Reston has never caused lethal infections in

humans. The discrepancies in pathogenicity among the Ebola virus subtypes are still largely misunderstood, as is the extreme virulence of Ebola Zaire (Takada and Kawaoka, 2001).

The clinical manifestations of Ebola virus infection are very severe and appear usually after a four to sixteen-day incubation period. Early symptoms include generalised aches and pains, fever, diarrhoea, nausea, vomiting and abdominal pains. After a week, these initial symptoms are followed by severe haemorrhage, multiorgan involvement and necrosis, generalised shock, and finally death (Baskerville et al., 1985; Fisher-Hoch et al., 1985). Ebola causes lesions in almost every organ, but most notably in the liver and spleen. One of the most striking features observed in fatal cases of the disease outbreaks in Gabon in 1996, is the massive intravascular apoptosis that develops at least 5 days before death (Baize et al., 1999). It has recently been suggested that in fatal cases, early interactions between Ebola virus and the immune system lead to defective immunity and massive apoptosis of lymphocytes and probably macrophages, whereas survivors display no alteration to their immune cells (Baize et al., 2000).

Recent epidemics have spurred an intensive search for an animal host that may act as a reservoir for Ebola virus. The high mortality rates in non-human primates suggest that these animals are unlikely to be the natural virus reservoir. Many animal species can be infected experimentally with Ebola virus, but detectable levels of the virus has not yet been detected in wild animals. Chronic infection of a mammal is nevertheless the most likely explanation for survival of the virus in nature (Zeller and Bouloy, 2000). In a hypothetical transmission cycle, bats are considered the most likely vertebrate hosts because they support replication of Ebola without causing disease (Monath, 1999). The major route of

Ebola virus transmission during outbreaks appears to be through direct contact with blood or infectious fluids from infected patients (Dowell et al., 1999). However, non-human primates have been experimentally infected by the aerosol, oral and conjunctival routes, suggesting that these means of transmission cannot be ruled out.

Detailed studies of Ebola have been difficult due to the infrequent and unpredictable outbreaks, the rapid course of infection, the remote areas where infections occur, and the requirement to study infectious materials in a biosafety level-4 laboratory. Analysis of patient cases and experimental infection of animals indicate that the rapid infection rate and extreme pathogenicity of Ebola infection allows little time for the production of effective anti-viral immune mechanisms in non-immune individuals. Current efforts are therefore being undertaken to develop effective anti-viral drugs and vaccines.

1.1.2 Search for vaccines

Putative targets for intervention by anti-viral drugs include inhibiting steps in viral replication or pathogenesis, as well as finding compounds that specifically induce components of the immune system, such as interferon (IFN). But so far, efforts to develop effective anti-viral drugs against Ebola have not yet been successful. Vaccination, thus, offers a promising intervention to prevent infection and limit the spread of the disease. In the case of Ebola virus, the conventional use of attenuated or inactivated virus preparations for the development of a vaccine is associated with serious safety risks (Wilson et al., 2001a). As a result, most of the vaccine approaches for Ebola have examined the ability of one or several Ebola proteins to serve as protective antigens. A number of vaccine strategies have been

tested so far: use of live vaccinia virus vectors, DNA vaccination, Venezuelan equine encephalitis virus (VEEV) replicons, and a DNA prime with adenovirus boost strategy.

Live vaccinia viruses have been used as expression vectors for a number of Ebola virus proteins, and their efficacy as vaccines were tested in a guinea-pig model (Gilligan, 1997) and in mice (Wilson et al., 2001b). In guinea pigs, only the glycoprotein (GP)-vaccinated animals exhibited any level of protection. More importantly, these studies highlighted the need to thoroughly examine the immune mechanisms mediating protection as well as the ability of vaccine vectors to effectively induce those responses. So far the different studies obtained variable results and it is not clear whether this is due to the different animal models examined or to the different vectors and vaccination schedules used in each study. In addition, a major disadvantage of this vector approach is that immunological competition may result from the production of the numerous vaccinia virus proteins in the host cell, thus lowering immune responses to the particular Ebola virus protein expressed by the vector.

In the case of VEEV, the vaccine is prepared by replacing the VEEV structural protein genes with a gene encoding an Ebola protein, resulting in a self-replicating RNA molecule (replicon) that can be packaged into a virus-like particle with helper RNAs encoding the VEEV structural proteins (Pushko et al., 1997). The efficacy of VEEV replicons has been tested both in mice and in guinea pigs, and so far the results are very encouraging (Pushko et al., 2000). The more interesting results came from the use of VEEV replicons expressing either the glycoprotein (GP) or the nucleoprotein (NP) of Ebola. For instance, in guinea pigs

vaccination with Ebola GP protected between 60 and 100% of the guinea pigs, depending on the strain.

DNA vaccination, as with the vaccinia and VEEV vectors, has been shown to influence both humoral and cellular immune activation pathways, but is not subjected to concerns about anti-vector immunity. DNA vaccines have been shown to elicit specific immune responses to Ebola virus antigens and to protect guinea pigs and mice against challenge with Ebola virus adapted to produce lethal infection in rodents (Vanderzanden et al., 1998; Xu et al., 1998). To obtain a more potent vaccine, Sullivan *et al.* have developed a new vaccine strategy (Sullivan et al., 2000). In a first step, the animals are vaccinated with DNA encoding Ebola virus proteins, and secondly, the immune response to these proteins is boosted with attenuated forms of the adenovirus expressing Ebola proteins. This strategy has been shown to be effective in protecting guinea pigs against Ebola infection and preliminary results show that primates can also be immunised against the lethal effects of Ebola virus. The use of a complex vaccine expressing Ebola structural proteins from diverse geographic isolates generated a strong antigen-specific immune response and resulted in the survival of all the immunised primates after challenge with a lethal dose of the most virulent Ebola virus, the Zaire strain.

A better understanding of the immune mechanisms involved in the protection against Ebola and the current development of a reverse genetics system to generate infectious Ebola virus should facilitate the development of efficacious vaccines or treatments against Ebola virus, which are suitable for use in humans.

1.2 General characteristics

1.2.1 Genome organisation and replication

The genome of filoviruses consists of a molecule of linear, non-segmented, negative-strand RNA, which is not polyadenylated. The genome (19,000 nucleotides) represents 1.1% of the total virion weight.

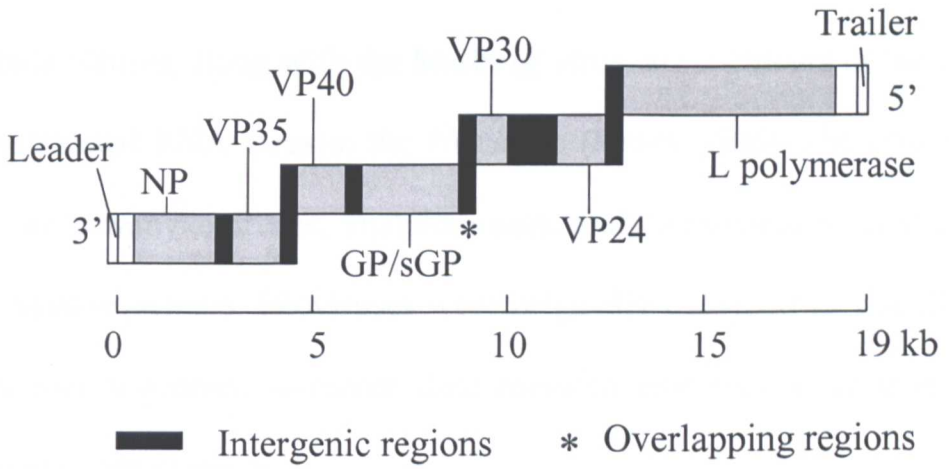

Figure 27: Ebola (Zaire) virus genome organisation.

The genome displays a linear arrangement (figure 27) of the genes, in the following order: 3′ untranslated region, nucleoprotein (NP), viral structural protein (VP) 35, VP40, glycoprotein (GP), VP30, VP24, L polymerase and finally a 5′ untranslated region. The genes are flanked at their 3′ and 5′ ends by highly conserved transcriptional start and termination sequences, respectively.

The genetics of filoviruses are most probably similar to those of rhabdo- and paramyxoviruses. Transcription and replication take place in the cytoplasm of infected cells. The 3′ untranslated region of the genome probably provides the entry site for the polymerase. Filovirus genomes are transcribed by the virion associated RNA-dependent RNA polymerase L, to yield monocistronic subgenomic mRNA species, which are complementary to viral genomic RNA. Transcription efficiency may be influenced by gene order, overlapping genes and

the presence of duplicated termination sites. Replication of the genome is achieved by the synthesis of a full-length complementary (positive sense) anti-genome, which in turn serves as the template for the synthesis of progeny negative-strand RNA (Peters, 1996).

1.2.2 Virion properties

Ebola viruses, along with the Marburg virus, are members of the family of negative-stranded RNA viruses, the *Filoviridae* (Peters, 1996). The filoviruses, in addition to paramyxoviruses, rhabdoviruses and bornaviruses, constitute the order of *Mononegavirales*. Filoviruses were originally classified as rhabdoviruses, but more recent genome sequence data revealed that they were more closely related to paramyxoviruses.

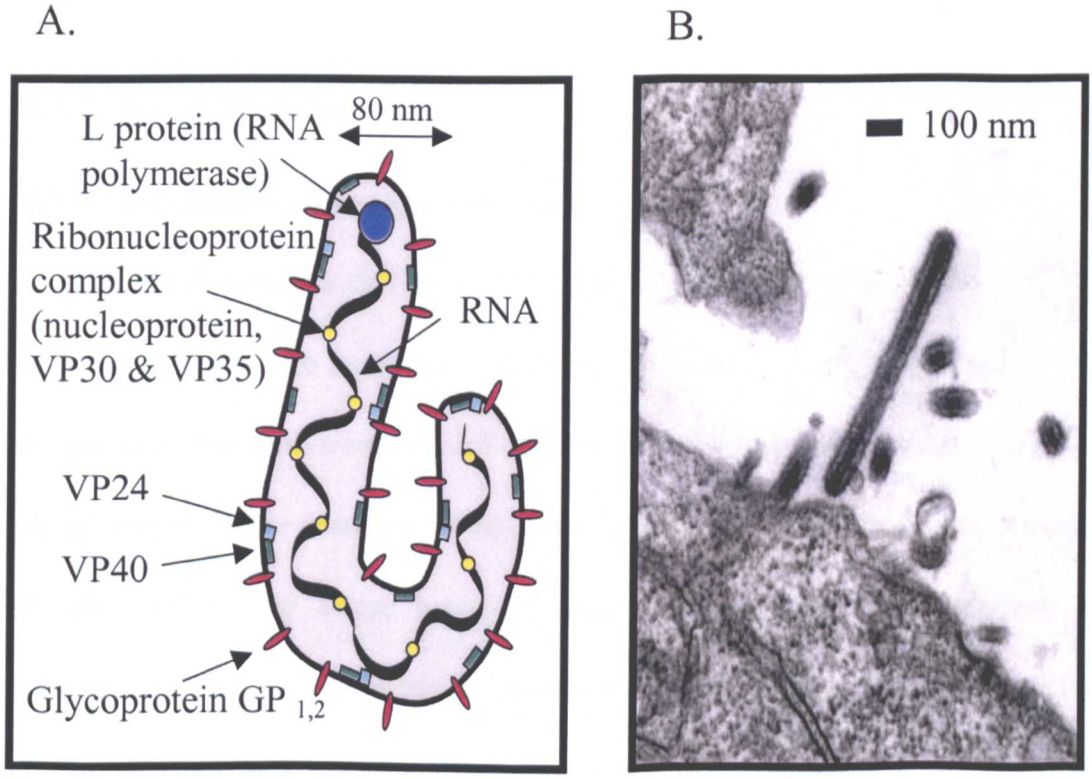

Figure 28: (A) Schematic diagram of an Ebola virus particle. (B) Electron microscopy image of Vero cells infected with Ebola virus http://www.pasteur.fr/recherche/unites/scme/portfolio/virus/Ebola3.html.

By electron microscopy (figure 28B), filovirus particles are found to be pleomorphic, appearing as long filamentous, occasionally branched forms, or as 'U'-shaped, 'b'-shaped or circular forms. The particles have a uniform diameter of 80 nm (figure 28A) (Geisbert and Jahrling, 1995).

Virions contain a nucleocapsid consisting of a dark, central space (20 nm in diameter) surrounded by a helical capsid (50 nm in diameter) bearing cross-striations with helix pitch of 5 nm. Within the nucleocapsid is an axial channel of 10-15 nm. A lipoprotein unit membrane envelope, derived from the host cell membrane, surrounds the helical nucleocapsid. Spikes of approximately 7 nm in length, spaced at approximately 10 nm intervals are visible on the virion surface. Virus particles have a density in potassium tartrate of 1.14 g/cm^3 (Peters, 1996).

1.3 Viral proteins

1.3.1 Importance of the glycoprotein

The RNA genome of Ebola virus encodes seven structural proteins (figure 28A). Four proteins are associated with the viral ribonucleoprotein complex (the nucleoprotein, the polymerase and the viral structural proteins 30 and 35), the single glycoprotein is inserted in the envelope, and the viral structural proteins VP40 and VP24 (the latter being unique to filoviruses) are believed to be membrane-associated. The glycoprotein forms the spikes (of approximately 7 nm in diameter) that are found on the surface of the virus (Peters, 1996), and is believed to have a central role in many steps during virus infection. The envelope glycoprotein consists of a trimeric type I transmembrane glycoprotein (Takada et al., 1997); its middle region is variable, extremely hydrophilic, and carries the bulk

of the glycosylation sites for *N*- and *O*-glycans that account for nearly one third of the molecular weight (Feldmann et al., 1999).

As the sole viral surface protein, the filovirus glycoprotein is thought to be responsible for receptor binding and fusion with the target membrane, allowing for entry of the virus. Indeed, a mutant vesicular stomatitis virus (VSV) expressing the Ebola virus glycoprotein instead of the VSV G gene, was able to infect primate cells more efficiently than avian, insect, and other mammalian cells, corresponding to the host range tropism of Ebola virus (Takada et al., 1997). The asialoglycoprotein receptor found exclusively in hepatocytes was initially identified as a receptor for Marburg virus (Becker et al., 1995). This receptor recognises glycoproteins displaying *N*-linked sugar chains with terminal galactose residues. However, as filoviruses are pantropic and cells lacking this receptor (such as human endothelial and epithelial cells) are susceptible to these viruses, other cellular receptors must exist. Integrins, especially the beta1 group have been suggested to interact with the Ebola virus glycoprotein and perhaps be involved in Ebola virus entry into cells (Takada et al., 2000). Although Marburg and Ebola glycoproteins appear to interact with target cells by distinct processes (Chan et al., 2000), the similarity of cell tropism and pathological features of infection between these viruses suggests that a common molecule may mediate their entry into cells. Recently Chan *et al.* (Chan et al., 2001) identified a possible cofactor for cellular entry of both Marburg and Ebola viruses. This cofactor was identified as being the folate receptor-α (FR-α), a widely expressed GPI-anchored cell-surface receptor. Interestingly, the broad expression pattern of FR-α correlates very well with the known features of the filovirus life cycle. However, not all the cell types that are naturally permissive to filoviruses express FR-α. Thus, at present it is thought that

the glycoprotein may be able to interact with a number of cell surface receptors, one of which may be FR-α. In a similar way to that observed with human immunodeficiency virus (HIV) (Berger et al., 1999), filovirus glycoproteins may require a family of receptors to support their complex life cycle in a broad range of cells. Filoviral particles enter host cells by endocytosis (Geisbert and Jahrling, 1995) and fusion of the viral membrane with the host most probably occurs in the low pH endosomal compartment.

During its maturation, the glycoprotein is synthesised as a precursor protein, which is subsequently cleaved by a furin protease (Volchkov et al., 1998a) into two subunits, GP_1 (140kDa) and GP_2 (26 kDa), which are linked by a disulphide bond to form the mature $GP_{1,\ 2}$ complex. Filoviruses, along with influenza and retroviruses, display a common requirement for proteolytic cleavage of a single precursor protein into distinct receptor-binding (GP_1) and membrane fusion domains (GP_2). However, in the particular case of filovirus membrane fusion, the mechanism leading to activation of the fusion peptide is not clear. The three-dimensional structure (Weissenhorn et al., 1998) of the membrane fusion subunit, GP_2, of the Ebola glycoprotein, reveals that the isolated fusion domain forms a rod-shaped coiled coil, which resembles the coiled coils found in the fusion domains of influenza and retroviruses (e.g. HIV gp41). These strong similarities, along with specific details obtained from the structure of GP_2, suggest that filovirus fusion proteins may also undergo considerable conformational changes to mediate fusion of the virus envelope with the target membrane. Interestingly, more recently it was shown that this coiled coil region of GP_2 is essential to confer infectivity to the mutant VSV expressing Ebola glycoprotein (Watanabe et al., 2000). These results clearly indicate that GP_2 is important for the

function of the glycoprotein and that peptides corresponding to this region could perhaps act as efficient antiviral agents.

In contrast to all other filoviral genes, including the GP gene of Marburg virus (Will et al., 1993), the organisation and transcription of the GP gene of Ebola virus is unusual, in that transcriptional editing is needed to express the envelope glycoprotein (Sanchez et al., 1996; Volchkov et al., 1995). Unedited mRNA derived from the same gene encodes a non-structural smaller glycoprotein (sGP), which is extensively secreted from infected cells as a homodimer (Sanchez et al., 1998; Volchkova et al., 1998). The secreted GP (50 to 70 kDa) shares the first 295 amino acids with the transmembrane GP, but contains an additional 69 residues at the C-terminus. A third mRNA of the GP gene encoding another small glycoprotein (ssGP) has also been identified (Volchkov et al., 1995). The two different forms, transmembrane GP and secreted GP, have recently been shown to promote disease progression by different mechanisms (Yang et al., 1998). The secreted form of GP binds to neutrophils to prevent early events in activation, and possibly to interfere with the inflammatory responses that may provide innate immunity to the virus (Yang et al., 1998). As a result, viral replication is facilitated. Transmembrane GP, on the other hand, attaches to the host cells (strong binding to endothelial cells) and mediates viral entry. Expression of GP *in vivo* causes endothelial cell injury and loss that compromises vascular integrity, similar to the damage provoked by Ebola virus infection (Yang et al., 2000). Moreover significant amounts of GP_1 are released from infected cells, but its binding specificity is not known (Feldmann et al., 1999). Thus, a single viral gene is responsible for both paralysis of the host inflammatory response and for damage to the vascular system, the consequences of which are rapid progression of the

infection and the characteristic development of haemorrhages (Klenk et al., 1998). Interestingly, however, the highly pathogenic Marburg virus does not produce sGP because its glycoprotein gene is organised differently, but does release GP_1 in similar amounts to Ebola. In addition, several highly pathogenic Ebola virus variants only secrete very low amounts of sGP (Volchkov et al., 1998b). Thus the significance of sGP and the released GP_1 in the pathogenesis of filoviruses remains to be clarified.

1.3.2 The matrix protein, VP40

It has been suggested that VP40 corresponds to the matrix protein of other non-segmented negative strand RNA viruses, due to its biochemical behaviour following non-ionic detergent treatment of virions (Elliott et al., 1985; Kiley et al., 1988), the localisation of the gene within the viral genome, and its abundance in viral particles (Geisbert and Jahrling, 1995). VP40 is indeed the most abundant protein in filovirus particles, and a common feature of envelope virus matrix proteins is their localisation to the inner face of the viral membrane (Garoff et al., 1998). The matrix proteins have been shown to play a central role in virus assembly and budding. They associate with cellular membranes and interact at the site of assembly with the cytoplasmic tails of glycoproteins (Mebatsion et al., 1999; Sanderson et al., 1994; Schnell et al., 1998), as well as with the ribonucleoprotein particle (RNP) (Garoff et al., 1998). Immunoelectron microscopy studies of filoviruses suggest that the matrix protein VP40 and the nucleoprotein are closely associated during viral morphogenesis (Geisbert and Jahrling, 1995).

A number of three-dimensional structures are now available for retrovirus matrix proteins and for the influenza virus matrix protein, M1 (Sha and Luo,

1997). Although the sequences of retrovirus matrix proteins differ significantly, their three-dimensional structures are quite similar (Conte and Matthews, 1998). Matrix proteins from simian immunodeficiency virus (SIV-1) and human immunodeficiency virus (HIV-1) crystallise as trimers (Hill et al., 1996; Rao et al., 1995) and display a large basic surface, as can be found in matrix proteins from human T-cell leukaemia virus type II (HTLV-II) (Christensen et al., 1996), Mason-Pfizer monkey virus (Conte et al., 1997) and bovine leukaemia virus (Matthews et al., 1996). In addition, it has been proposed that influenza virus matrix protein M1 also has a basic region, which like in the cases of retrovirus matrix proteins, is suggested to mediate binding to the lipid bilayer (Ruigrok et al., 2000a). The retroviral matrix proteins are predominantly composed of α-helices, which are closely packed and joined by loops or regions of extended structure (Conte and Matthews, 1998). More recently, the three-dimensional structure of Ebola virus VP40 was solved (figure 29), corresponding to the first matrix protein structure from a negative-stranded non-segmented RNA virus (Dessen et al., 2000b). The VP40 monomer is an elongated, two-domain assembly, with dimensions of 40x50x25 Å. The N-terminal domain is folded into a β-sandwich consisting of six antiparallel strands arranged into two three-stranded β-sheets (β1-β6), while the C-terminal domain consists of one antiparallel triple-stranded β-sheet and an opposite set of three β-strands, forming a significantly bent β-sheet (β7-β12). The domains are connected by a flexible linker, which was not visible in the crystal structure. A search for structurally homologous proteins yielded no statistically significant matches, while superposition of the two domains revealed that they were topologically very similar, suggesting that the two domains may have arisen from a common ancestor by gene duplication. The structure of Ebola virus VP40

displays no resemblance to structures of retroviral matrix proteins, a finding that may be attributed to significant differences in the assembly and maturation of these viruses. In retroviruses, the matrix protein surrounds a spherical nucleocapsid, which originates from a polyprotein and mediates budding from the basolateral side of polarised cells, while the matrix protein of negative-stranded non-segmented RNA viruses is not processed by proteolytic cleavage, encloses a helical nucleocapsid and often promotes budding from the apical side of polarised cells (Garoff et al., 1998).

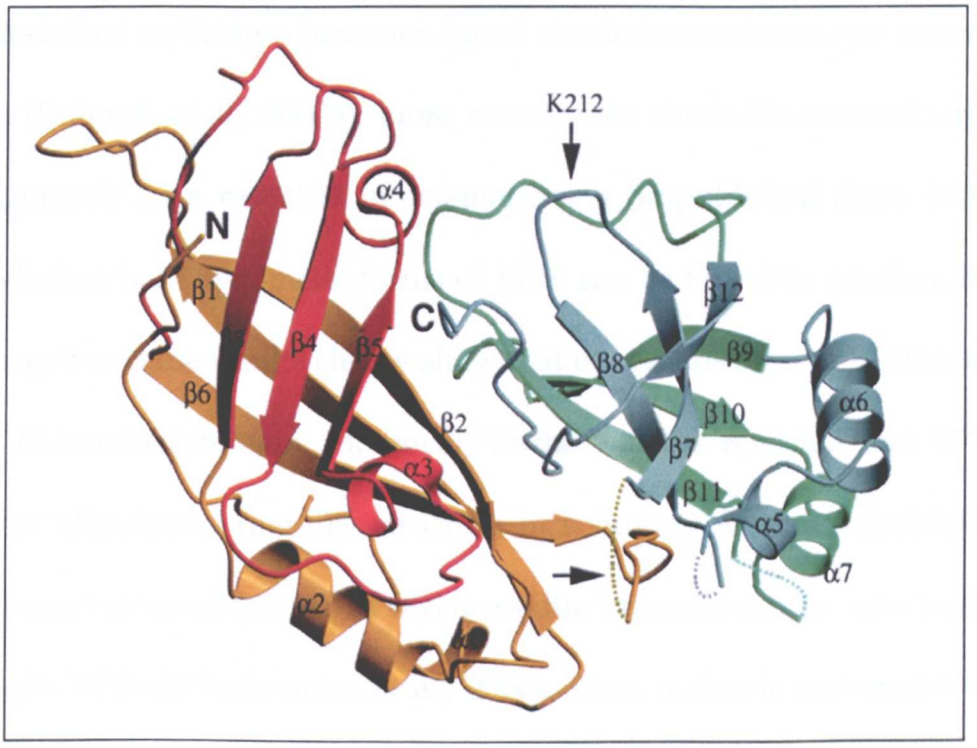

Figure 29: Ribbon diagram of Ebola virus VP40. The N-terminal domain is shown in orange and the C-terminal domain in blue. An arrow points to the trypsin cleavage site (after lysine 212).

From the structure, it becomes apparent that the VP40 domains are not tightly packed against each other. There are only very few interdomain side chain interactions and most of the contacts are made via a hydrophobic zipper region. The hydrophobic residues involved in these interdomain contacts are highly

conserved between Ebola and Marburg virus VP40 sequences. Moreover, trypsin cleaves full-length VP40 after lysine 212 (Ruigrok et al., 2000b), located between β-strands 7 and 8 in the C-terminal domain (figure 29), leading to the spontaneous dissociation of the C-terminal domain in solution. This behaviour further underlines the weak interaction between the N- and the C-terminal domains.

Interestingly, the removal of the C-terminal domain also results in the formation of an oligomeric, ring-like structure of Ebola VP40 (Ruigrok et al., 2000b; Scianimanico et al., 2000). Although, this ring-like structure had originally been described as being a hexamer, based on electron microscopy reconstruction studies (Ruigrok et al., 2000b), more recently the three-dimensional structure of the oligomeric form revealed an octameric ring (unpublished data). It is not yet clear whether both oligomeric forms of VP40 can be found in solution. However, taken together these results clearly show that oligomerisation of VP40 is mediated by the N-terminal domain. Moreover, further studies revealed that VP40 could adopt its oligomeric, ring-like structure in solution, by either destabilising the inter-domain interactions with a seven-residue truncation or by urea treatment of full-length VP40 (Scianimanico et al., 2000). Thus, a simple movement of the C-terminal domain appears to allow the N-terminal domain to form higher order oligomeric structures, which may be functionally relevant.

Biochemical characterisation of Ebola VP40 also revealed that, like other viral matrix proteins, VP40 could bind negatively charged lipid bilayers *in vitro* (Ruigrok et al., 2000b). The membrane binding capacity of viral matrix proteins has been described for a number of viruses: HIV-1 (Gottlinger et al., 1989; Zhou and Resh, 1996), vesicular stomatitis virus (VSV) (Chong and Rose, 1993; Zakowski et al., 1981), Sendai virus (Sanderson et al., 1993; Sanderson et al., 1994;

Stricker et al., 1994) and influenza virus (Enami and Enami, 1996; Kretzschmar et al., 1996; Ruigrok et al., 2000a; Zhang and Lamb, 1996). The binding of VP40 to liposomes is dependent on the presence of negatively charged phospholipids in the vesicles and could be abolished by adding 1M NaCl to the liposome/protein mixture, strongly suggesting that the electrostatic interactions play an important role in the binding (Ruigrok et al., 2000b). Further studies have now shown that the membrane binding capacity of Ebola VP40 is associated with the C-terminal domain, since removal of most of the C-terminal domain by trypsin digestion abolished the binding of VP40 (Ruigrok et al., 2000b). In addition, membrane association of VP40 has been shown to trigger a conformational change *in vitro,* resulting in the formation of the oligomeric ring-like structure associated with liposomes (Scianimanico et al., 2000). Viral matrix proteins from members of the Mononegavirales order have been shown to polymerise *in vitro* (Gaudin et al., 1997; Heggeness et al., 1982; McCreedy et al., 1990) and *in vivo* with the observation of crystalline lattices on the plasma membrane (Buechi and Bachi, 1982). Interestingly, this feature is also shared by retroviral Gag precursor proteins, which assemble into hexagonal arrays on membranes *in vitro* (Barklis et al., 1998). There is now growing evidence that a number of matrix proteins alone constitute a minimal assembly and budding machinery. Several matrix proteins have been shown to provoke the release of virus-like particles from cells expressing the matrix proteins by themselves. This is the case for the matrix protein from VSV (Justice et al., 1995; Li et al., 1993; Sakaguchi et al., 1999), from human parainfluenza virus type 1 (Coronel et al., 1999) and SIV-1 (Gonzalez et al., 1993). There is also growing evidence, however, that the concerted interaction of other viral proteins, such as the cytoplasmic tails of glycoproteins (Jin et al., 1997;

Sanderson et al., 1994; Schmitt et al., 1999; Schnell et al., 1998), the RNP particle (Garoff et al., 1998) and cellular factors (Garnier et al., 1996; Harty et al., 1999; Yasuda and Hunter, 1998), influences the efficiency of this particle formation. Interestingly, Ebola virus VP40 protein possesses two protein-protein interaction motifs at its N-terminus (figure 30): a PPxY motif, a putative WW domain binding motif (Bork and Sudol, 1994; Chen and Sudol, 1995; Sudol and Hunter, 2000), and a P(T/S)AP motif (Gottlinger et al., 1991), which has been proposed to interact with SH3 domains.

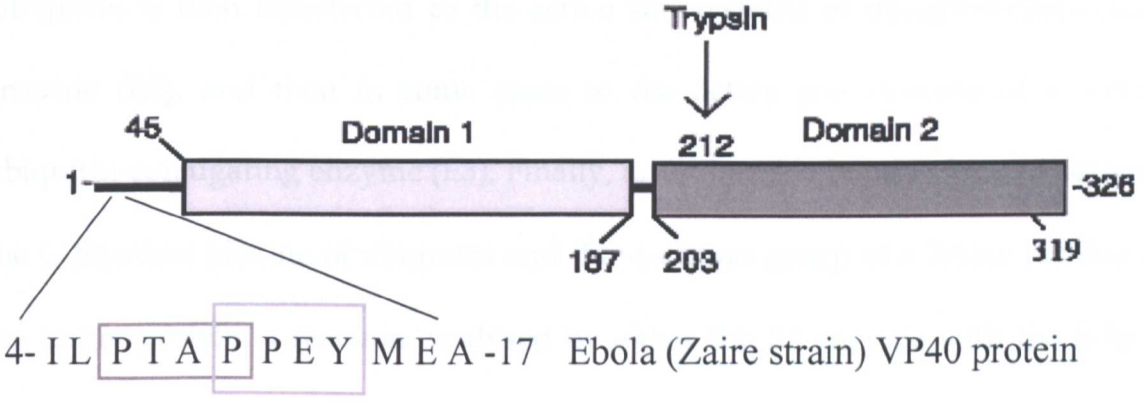

Figure 30: Schematic diagram of the domain organisation of Ebola VP40, showing the presence of the overlapping PTAP and PPEY motifs at the N-terminus.

2. Involvement of cellular machinery in virus assembly and budding

2.1 Ubiquitination

2.1.1 Roles of ubiquitination

Ubiquitination of proteins is a posttranslational modification in which a highly conserved 76-amino acid polypeptide, found in all eukaryotes, is attached to proteins. This process involves the sequential action of three enzymes (figure 31): E1, E2 and E3. The E1 enzyme (ubiquitin-activating) first activates ubiquitin in

an ATP-dependent reaction by forming a thioester bond at its active-site cysteine with the C-terminus of ubiquitin.

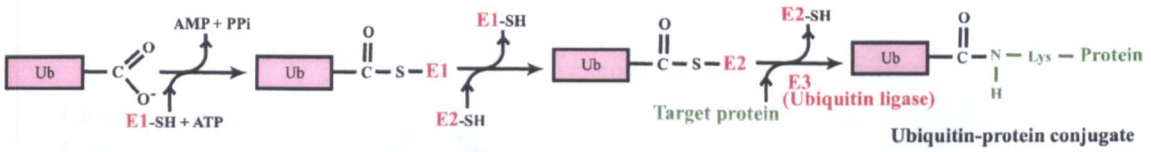

Figure 31: Schematic diagram of the ubiquitination pathway, involving three enzymes: E1, E2 and E3.

Ubiquitin is then transferred to the active site cysteine of ubiquitin-conjugating enzyme (E2), and then in some cases to the active site cysteine of a second ubiquitin-conjugating enzyme (E3). Finally, an isopeptide bond is formed between the C-terminal glycine of ubiquitin and the ∈-amino group of a lysine residue on the target protein, a reaction catalysed by either the E2 enzyme with the help of the E3 enzyme, or directly by the E3 enzyme. Polyubiquitin chains are formed on proteins by the conjugation of additional ubiquitin moieties to one of several lysine residues in the ubiquitin molecules previously attached to the protein (Chau et al., 1989).

Over the past few years it has become apparent that ubiquitination is involved in a large range of cellular processes (reviewed in (Pickart, 2001)). The most studied role of ubiquitination is its involvement in targeting proteins for degradation by the 26S proteasome (Hochstrasser, 1996; Voges et al., 1999). This process involves the polyubiquitination of the proteins to be degraded. More recently, however, other functions for ubiquitin are being discovered, most of which involve monoubiquitination of proteins (Hicke, 2001). Monoubiquitination appears to be involved in a number of intracellular sorting pathways (Beck et al.,

1999) and endocytosis of cell-surface receptors (Hicke, 1997; Hicke, 1999; Strous and Govers, 1999) (figure 32).

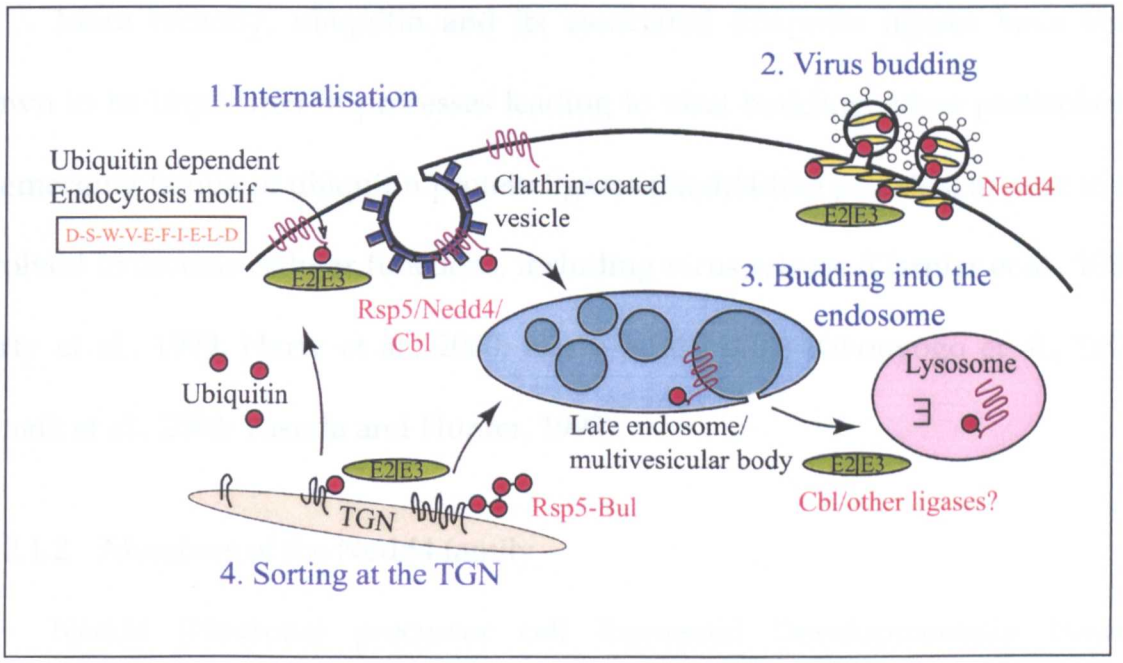

Figure 32: Schematic representation of the various steps at which ubiquitin is required.

Ubiquitination is required for cargo sorting in at least three distinct protein transport events: at the plasma membrane, at the late endosome, and at the *trans*-Golgi network (TGN). Recent studies have provided the first hints as to how ubiquitin is recognised as a sorting signal, by ubiquitin binding proteins (Hofmann and Falquet, 2001). In this work they identified a 20 amino acid sequence, called the ubiquitin-interacting motif (UIM) that is present in a number of proteins from yeast to humans. Several of these newly identified ubiquitin-interacting proteins are known components of the endocytic pathway. At the late endosome, there is another candidate for an ubiquitin binding protein involved in multivesicular body (MVB) sorting: Vps23 and its mammalian homologue Tsg101 (tumour susceptibility gene) (Katzmann et al., 2001). Vps23 is part of a protein

complex that binds ubiquitin *in vitro* and to ubiquitinated MVB cargo *in vivo* (Dupre et al., 2001).

More recently, ubiquitin and its associated ubiquitin ligases have been shown to be implicated in processes leading to viral budding. More particularly, an emerging family of ubiquitin protein ligases (Nedd4-like proteins) appear to be involved in diverse cellular functions, including virus release (Garnier et al., 1996; Harty et al., 1999; Harty et al., 2000; Harty et al., 2001; Kikonyogo et al., 2001; Patnaik et al., 2000; Yasuda and Hunter, 1998).

2.1.2 Members of the Nedd4 family

Nedd4 (Neuronal precursor cell Expressed Developmentally Down-regulated gene 4) was originally identified as a developmentally regulated mouse gene, highly expressed in early embryonic central nervous system (Kumar et al., 1992). Further studies revealed that Nedd4 was actually expressed at varying levels in different tissues (Kumar et al., 1997). The characteristic feature of the Nedd4 family of ubiquitin ligase (E3) enzymes is its domain organisation: a C2 domain, 2 to 4 copies of WW domains and a HECT (homologous to E6-AP C-terminal) domain, which is the ubiquitin-protein ligase domain (Harvey and Kumar, 1999). The C2 domain is believed to regulate the function of proteins by mediating their translocation to phospholipid membranes in response to an increase in intracellular Ca^{2+} concentration (Rizo and Sudhof, 1998). WW domains derive their name from the presence of 2 highly conserved tryptophan residues and a conserved proline residue in a sequence of 35-40 amino acids (Sudol, 1996). These domains consist of a hydrophobic core surrounded by β-sheets containing a number of charged residues (Macias et al., 1996), and have a preference for

binding small proline-rich sequences: PPxY or more rarely, PPLP motifs. Finally, the large HECT domain (~ 350 residues) is found at the C-terminus of Nedd4 molecules and comprises the substrate-specificity arm of the ubiquitin pathway (Hershko and Ciechanover, 1998). HECT domain proteins are a major subclass of E3 enzymes and contain a conserved cysteine residue that is capable of forming a thioester bond with ubiquitin (Huibregtse et al., 1995). To date, Nedd4 orthologues have been found in yeast, mouse, rat and human and the presence of multiple Nedd4-like proteins in humans (14 identified cDNA sequences) suggests that many mammalian proteins are likely to be modified through ubiquitination by Nedd4 family members.

At present, clear functional evidence is only available for mammalian Nedd4 and its yeast homologues Rsp5p/Npi1p (*S. cerevisiae*) and Pub1p (*S. pombe*). There is growing evidence for a role of mammalian Nedd4 in the regulation of epithelial sodium channels (Goulet et al., 1998; Harvey et al., 1999). The WW domains of Nedd4 can interact with the PPxY motifs in the epithelial Na^+ channel subunits (α, β, and γ) (Farr et al., 2000; Staub et al., 1996) and the α and γ subunits are ubiquitinated (Staub et al., 1997). Moreover, this interaction leads to channel inhibition. Interestingly, the PPxY motif of either the β or the γ subunit is disrupted in Liddle's syndrome (increased blood pressure), leading to increased Na^+ current, in part due to an increase in the number of Na^+ channels at the cell surface. It has, therefore, been hypothesised that failure of Nedd4 to bind to and ubiquitinate Na^+-channel subunits, prevents channels from being tagged for degradation, thus resulting in an accumulation of channels at the plasma membrane (Goulet et al., 1998; Snyder et al., 2001). The yeast homologue of Nedd4, Rsp5p, is an essential protein implicated in a number of cellular processes

regulated by the ubiquitin system, including endocytosis (Harvey et al., 1999; Rotin et al., 2000). Evidence for a role in endocytosis came from the study of an *rsp5/npi1* mutant in which the ubiquitination and endocytosis of the uracil and general amino acid permeases that occur in response to changes in nutrient availability and stress were impaired (Galan et al., 1996; Springael and Andre, 1998). Interestingly, the multiple domains of Rsp5, including the three WW domains, each have an important role although at different steps in the endocytosis process (Dunn and Hicke, 2001; Gajewska et al., 2001; Springael et al., 1999; Wang et al., 1999; Wang et al., 2001). Rsp5/Npi1p is also believed to play a role in minichromosome maintenance (Yashiroda et al., 1996), mitochondrial/cytoplasmic protein distribution (Zoladek et al., 1997) and is required for vegetative growth, sporulation and the stress response (Kanda, 1996). In addition, human Nedd4 and Rsp5p also potentiate hormone-dependent activation of transcription by progesterone and glucocorticoid receptors, in a manner apparently independent of ubiquitin-protein ligase function and interactions through the WW domains (Imhof and McDonnell, 1996).

Several structures of E3 enzymes have been solved in complex with either their E2 enzyme (E6AP/UbcH7 complex; (Huang et al., 1999)) or the binding partner of the WW domains (Nedd4 WW domain-ENaC complex; (Kanelis et al., 2001)). The E6AP ubiquitin-protein ligase (E3) is a member of the HECT E3 class of enzymes and displays approximately 30% sequence identity with the HECT domains of yeast Rsp5p and human Nedd4. The E6AP/UbcH7 complex has a U-shaped structure (Huang et al., 1999). The HECT domain consists of two lobes that pack loosely across a small interface and are connected by a three-residue hinge. The large amino-terminal lobe has a mostly α-helical structure with an elongated

shape, while the smaller carboxyl-terminal lobe has an α/β structure. The catalytic cysteine residue is found in a broad cleft at the junction between the two lobes. UbcH7 (E2 enzyme) binds in a large hydrophobic groove on the amino-terminal lobe of E6AP. The E2-binding groove consists of residues that are only moderately conserved but maintain their hydrophobic character in the HECT domains of Rsp5p and Nedd4 (Huang et al., 1999). In addition, the solution structure of rat Nedd4 WW domain 3 and a peptide containing the PPxY motif derived from the β subunit of the epithelial sodium channel (ENaC) was recently solved (Kanelis et al., 2001). The structure of the WW domain in the complex is very similar to known structures of other WW domains (Huang et al., 2000; Macias et al., 1996; Macias et al., 2000; Verdecia et al., 2000). The peptide-binding site located on one face of the sheet and consisting of a groove nearly orthogonal to the β-strands, is formed by residues throughout the WW domain. Binding of the peptide is mediated primarily by hydrophobic groups (Kanelis et al., 2001).

2.2 Evidence for virus-cell interactions

2.2.1 Ubiquitin and virus budding

The first evidence for the implication of ubiquitin in virus budding came from the observation that avian retroviruses contain unexpectedly large amounts of free ubiquitin, reaching levels 5-fold higher than that of free ubiquitin in the cytosol (Putterman et al., 1990). Similar amounts of ubiquitin have more recently been found in HIV-1, SIV, and murine leukaemia virus (MLV) (Ott et al., 1998), and some of this ubiquitin appears to be conjugated to the viral Gag proteins. The retrovirus Gag proteins are functionally equivalent to the matrix proteins of negative-strand RNA viruses. Like matrix proteins, Gag proteins can associate

with the cellular membrane and bud from cells independently of other viral proteins (Wills et al., 1994), a phenomenon dependent on having a functional L-domain. These L domains are believed to recruit the cellular machinery needed for virus-cell separation on the plasma membrane by protein-protein interactions with specific proline-rich motifs: PPxY and/or P(T/S)AP motifs. At present, the PPxY motifs within RSV Gag, VSV M, rabies virus M, Ebola VP40 proteins have been shown to interact with WW domains of specific cellular proteins, including members of the Nedd4 family (Harty et al., 1999; Harty et al., 2000; Harty et al., 2001; Kikonyogo et al., 2001). In addition the latent membrane protein 2A of Epstein-Barr virus also possesses several PPxY motifs that are able to interact specifically with members of the Nedd4-like ubiquitin-protein ligase family (Ikeda et al., 2000; Winberg et al., 2000).

Over the past few years, evidence for the implication of ubiquitination in viral budding has accumulated. Notably, inhibition of the proteasome, which reduces the level of free ubiquitin in the cytosol, interferes with the release of virus-like particles in rhabdo- and retroviruses (Harty et al., 2001; Patnaik et al., 2000; Schubert et al., 2000). Moreover, a functional RSV L domain and other unrelated L domains (from HIV-1, HTLV-1 and Ebola virus) induce the ubiquitination of RSV Gag protein (Strack et al., 2000). Nevertheless, the mechanism by which cellular factors enhance the budding process of viruses remains unclear.

2.2.2 Recent involvement of the vacuolar protein sorting pathway

Recent data also suggests that the mammalian homologue of Vps23, Tsg101, is implicated in virus budding (VerPlank et al., 2001). Tsg101 has been

found to bind to the P(T/S)AP motif within HIV-1 p6 (Garrus et al., 2001; VerPlank et al., 2001) and to facilitate viral budding. Point mutations within this P(T/S)AP motif arrest viral release at a very late stage (Gottlinger et al., 1991; Huang et al., 1995). Tsg101 was originally identified as a cellular transforming gene in a genetic screen (Li and Cohen, 1996), and shares homology with several inactive ubiquitin-conjugating enzymes identified in yeast as being important for late endosomal trafficking. Mutant Tsg101 cells appear to be defective in the delivery of proteins to late endosomal compartments, resulting in recycling of these proteins back to the cell surface (Babst et al., 2000). Tsg101 positively regulates lysosomal degradation of some cell surface proteins (Li et al., 2001). Tsg101 and its yeast counterpart, Vps23, are unlikely to catalyse ubiquitination because they lack the active-site cysteine, but it has been suggested that they may be involved in ubiquitin recognition or modification. Tsg101 may bind to ubiquitinated membrane proteins to direct them into forming multivesicular late endosomes, or may act as a regulator of E3-dependent ubiquitination of the cargo (Lemmon and Traub, 2000). The finding that Tsg101 is also involved in virus budding strongly suggests that the ubiquitination and vacuolar protein sorting machinery (Perez and Nolan, 2001) may be required for viral budding, a process that is topologically very similar to the formation of multivesicular bodies.

3. Aims of the project

Because of their extreme pathogenicity, studying filoviruses has been a very difficult task and our understanding of the molecular mechanisms underlying the virus life cycle is very limited. The matrix protein, VP40, of Ebola virus has been the focus of many biochemical studies, most of which were carried out in *in vitro*

systems. In this work, we investigated in a cellular system, the role of VP40 in the final steps of virus replication: the exit from the cell. Matrix proteins of rhabdoviruses and retroviruses have been shown to elicit the release of virus-like particles when transfected into cells by themselves. It was, therefore, of much interest to see whether Ebola VP40, alone, could provoke the budding process and whether the efficiency of this release was dependent on a particular amino acid sequence. In addition, the growing evidence for the involvement of the ubiquitination and vacuolar protein sorting pathways in virus assembly and budding, pressed us to look at the possible interactions of VP40 with a member of the human Nedd4 family of ubiquitin ligases and human Tsg101. Do very different viruses subvert the same cellular protein machinery to exit infected cells?

Chapter 5: Results

1. Vesicular release of Ebola VP40

1.1 Expression constructs for mammalian cell expression of VP40

Ebola virus VP40 consists of two domains that form a closed conformation in solution (figure 33). The N-terminal part, which is not present in the crystal structure (Dessen et al., 2000b), contains a PPxY motif and a P(T/S)AP motif. cDNAs corresponding to VP40 (1-326), VP40 (31-326) and a C-terminally truncated form, VP40 (31-212) (figure 33) were cloned into the pcDNA 3.1 vector (Invitrogen) for expression in mammalian cells.

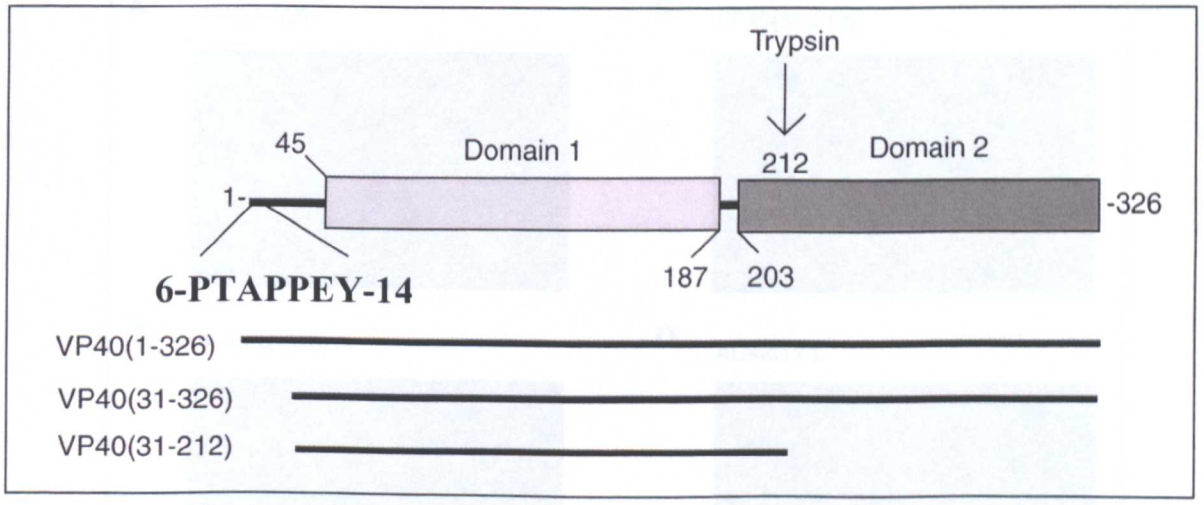

Figure 33: Schematic representation of the two structurally related domains of VP40. VP40 constructs used for mammalian cell expression are shown below.

VP40 (31-326) is missing the PPxY and P(T/S)AP motifs at the extreme amino-terminus of VP40. The three cDNAs were transfected into human

embryonic kidney 293T cells and expression was detected by either indirect immunofluorescence or by Western Blotting, using a VP40-specific anti-serum.

1.2 Cellular localisation of VP40

To study the cellular localisation of VP40 in transfected 293T cells, indirect immunofluorescence staining of the cells was carried out 72 hours post-transfection. Both VP40 (1-326) and VP40 (31-326) are expressed throughout the cytoplasm but concentrate in patches along the plasma membrane as seen by indirect immunofluorescence (figure 34A and B). In contrast, the C-terminally truncated construct, VP40 (31-212), does not localise to the plasma membrane and shows only diffuse distribution throughout the cytoplasm (figure 34C). No staining can be seen with non-transfected or mock-transfected cells (data not shown).

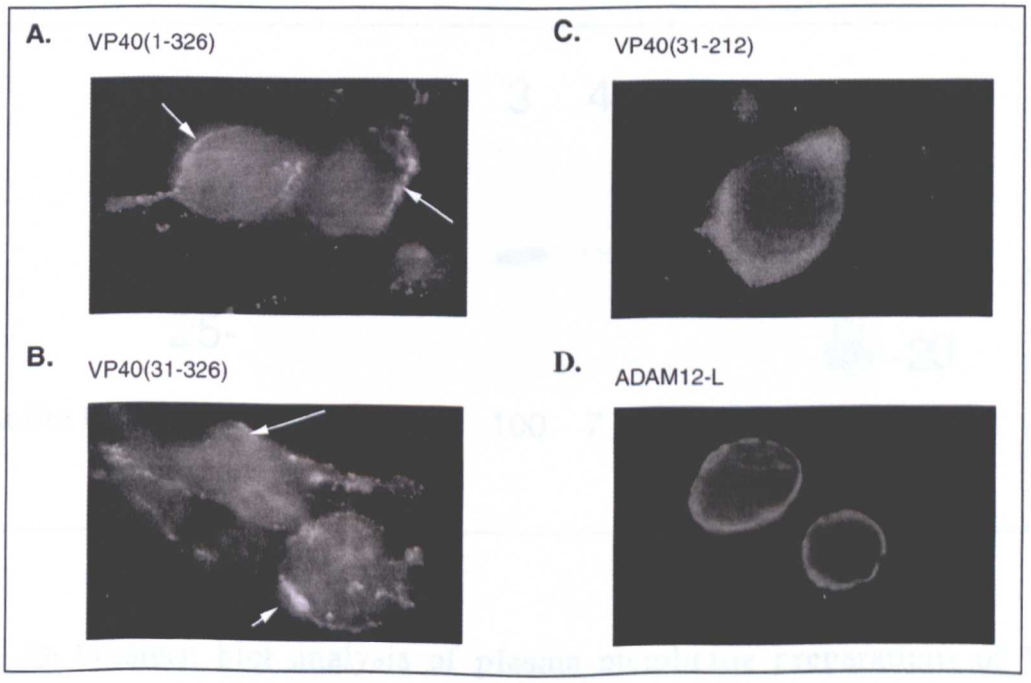

Figure 34: Intracellular distribution of VP40 constructs and ADAM 12-L as determined by indirect immunofluorescence, using a rabbit polyclonal antiserum to detect VP40 and an anti-*myc* antibody to visualise ADAM 12-L.

These results were confirmed by the analysis of plasma membrane preparations derived from transfected 293T cells expressing VP40 (1-326), VP40 (31-326) and VP40 (31-212). Full-length VP40 migrates at approximately 37 kDa, the N-terminal truncated form slightly faster and C-terminally truncated VP40 migrates at approximately 25 kDa when analysed by SDS-PAGE and western blotting (figure 35). Both constructs, VP40 (1-326) and VP40 (31-326) associate with the plasma membrane (Figure 35). Comparison of the total protein expressed (Figure 35, lanes 1 and 3) and the amount found to interact with the plasma membrane shows approximately 7 % membrane association for both constructs (Figure 35, lanes 2 and 4). This indicates that the first 31 residues of VP40 are dispensable for its cellular localisation and does not affect the efficiency of membrane association, similar to the results obtained *in vitro* (Ruigrok et al., 2000b).

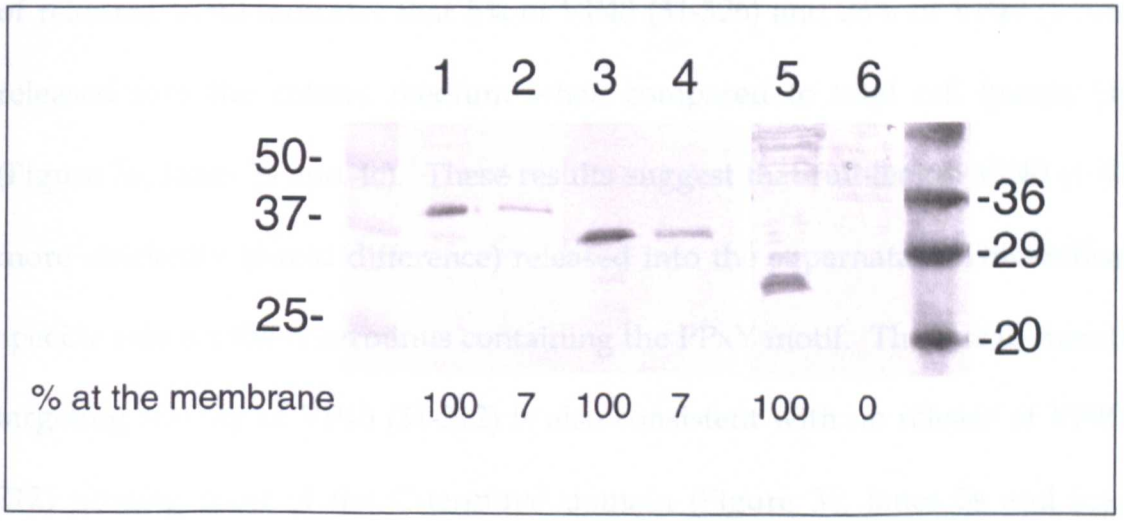

Figure 35: Western blot analysis of plasma membrane preparations of VP40 expressed in 293T cells. Lanes 1, 3 and 5 correspond to the total amount of VP40 (1-326), VP40 (31-326) and VP40 (31-212) respectively, expressed from total cell lysates ($2x10^5$ cells). Lanes 2, 4 and 6 correspond to the amount of VP40 (1-326),

VP40 (31-326) and VP40 (31-212) respectively, found to be associated with the plasma membrane.

Deletion of most of the C-terminal domain of VP40 resulted in solely cytoplasmic expression of VP40 (31-212) and no protein was detected to be associated with the plasma membrane (figure 35, lanes 5 and 6), which confirms the role of the C-terminal domain in membrane targeting *in vivo*. This is also consistent with the immunofluorescence staining (figure 34C) and liposome-binding assays performed *in vitro* (Ruigrok et al., 2000b).

1.3 Release of VP40 into the cell culture supernatant

Western blot analysis of the expression pattern of all three constructs analysed indicates that full-length VP40 (1-326) as well as VP40 (31-326) are both released into the cell culture medium (Figure 36, lanes 1s and 3s). Quantification of released VP40 indicates that 5% of VP40 (31-326) and 23% of VP40 (1-326) are released into the culture medium when compared to total cell lysates (100%) (Figure 36, lanes 3s and 4c). These results suggest that full-length VP40 (1-326) is more efficiently (5-fold difference) released into the supernatant. This indicates a specific role for the N-terminus containing the PPxY motif. The loss of membrane targeting activity of VP40 (31-212) is also consistent with no release of VP40 (31-212) missing most of the C-terminal domain (Figure 36, lanes 5s and 6c). To further substantiate that the presence of VP40 in the supernatant is not due to the release of protein by cell lysis, 293T cells were also transfected with a full-length clone of ADAM 12-L, a type I membrane protein (Gilpin et al., 1998). Transfections of all four constructs were done in parallel, and no significant cell death was observed at the time harvesting the cells and the supernatants. ADAM

12-L is clearly expressed at the cell surface of 293T cells as expected (figure 34D) and an expression band corresponding to the correct molecular weight can be seen in a whole cell extract but not in the cell culture supernatant (figure 36, lanes 7s and 8c).

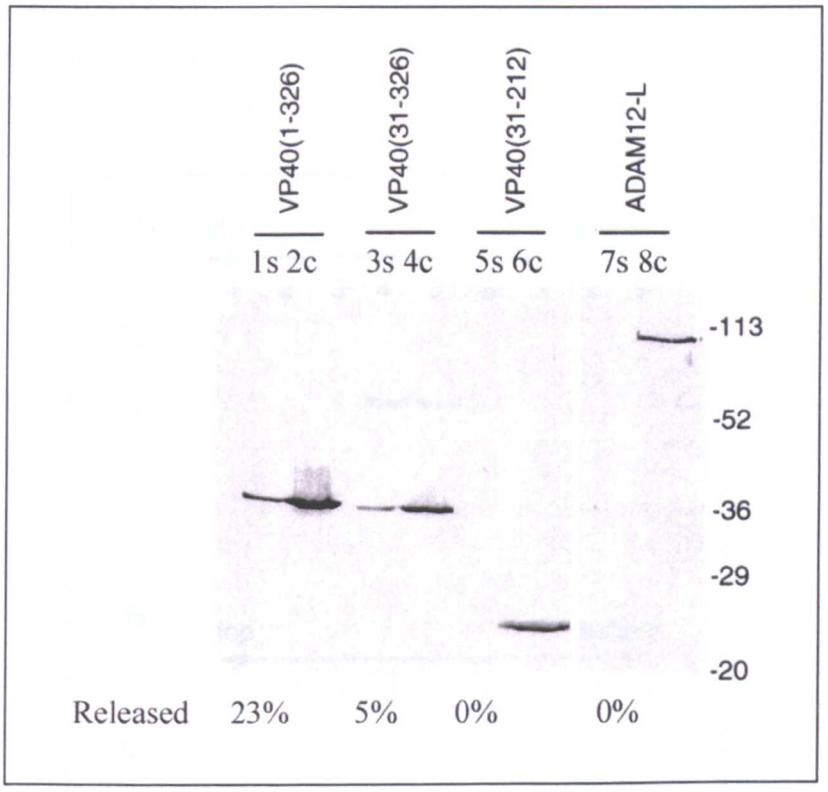

Figure 36: Western blot analysis of VP40 expression in the cell extracts (c) and supernatants (s) of transfected 293T cells. 293T cells were transfected with VP40 (1-326) in lanes 1 and 2, with VP40 (31-326) in lanes 3 and 4, with VP40 (31-212) in lanes 5 and 6 and with human ADAM 12-L in lanes 7 and 8. VP40 (1-326) was found to be released into the supernatant 4-5 fold more efficiently than VP40 (31-326).

These results suggest that specific membrane association of VP40 (1-326) and VP40 (31-326) is necessary for the release of VP40 into the medium. This interaction does not occur with VP40 expressed only in the cytoplasm (such as VP40 (31-212)) or with a plasma membrane anchored protein such as ADAM 12-L, as expected.

The supernatants of cells transfected with VP40 constructs were further analysed by sucrose gradient centrifugation. These results show that VP40 (1-326) (figure 37A) as well as VP40 (31-326) (figure 37B) migrate towards the top in a discontinuous sucrose gradient and most of the protein is found in fractions 4 to 6 or 3 to 5 (figure 37A and B).

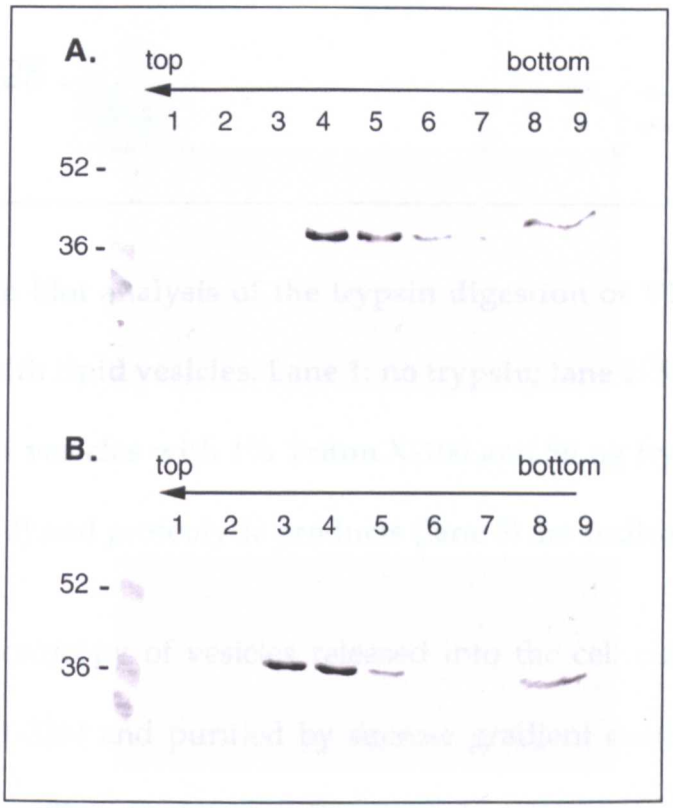

Figure 37: Western blot analysis of the sucrose gradient floatation assay with (A) VP40 (1-326) and (B) VP40 (31-326) released into the cell culture medium.

Some protein remains in the bottom fractions 8 and 9 (figure 37A and B). Therefore, both VP40 proteins probably float with vesicles in upper fractions containing 20 to 30 % sucrose. Trypsin digestion of vesicle-bound VP40 (1-326) also indicates that it is protected from proteolysis. Incubation of VP40 (1-326) derived from fraction 4 of the sucrose gradient (figure 37) with trypsin showed no proteolysis (figure 38; lane 2). However, solubilisation of VP40 from the same sucrose gradient fraction with Triton X100 shows that detergent treatment renders

VP40 sensitive to trypsin, resulting in smaller proteolytic products (figure 38, lane 3).

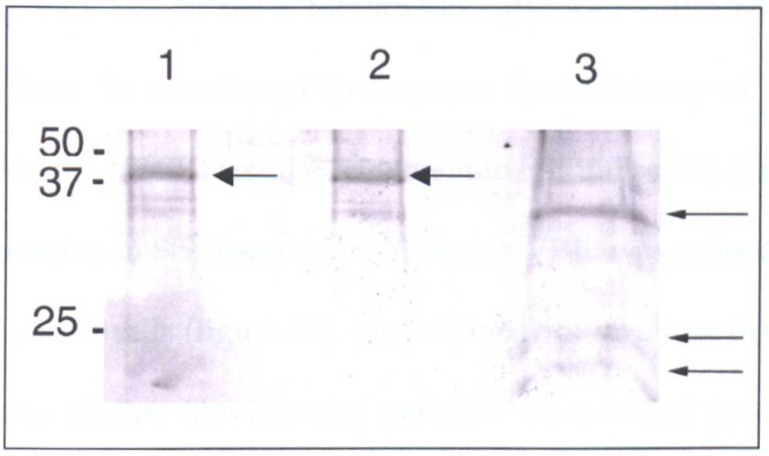

Figure 38: Western blot analysis of the trypsin digestion of VP40 (1-326) found to be associated with lipid vesicles. Lane 1: no trypsin; lane 2: 50 ng trypsin; lane 3: solubilisation of vesicles with 1% Triton X-100 and 50 ng trypsin. Full-length VP40 (lanes 1 and 2) and proteolytic products (lane 3) are indicated by arrows.

Electron microscopy of vesicles released into the cell culture supernatant containing VP40 (1-326) and purified by sucrose gradient centrifugation reveals the presence of rare filamentous virus-like particles. A typical particle shows a diameter of approximately 80 nm and a length of 1000 nm (figure 39).

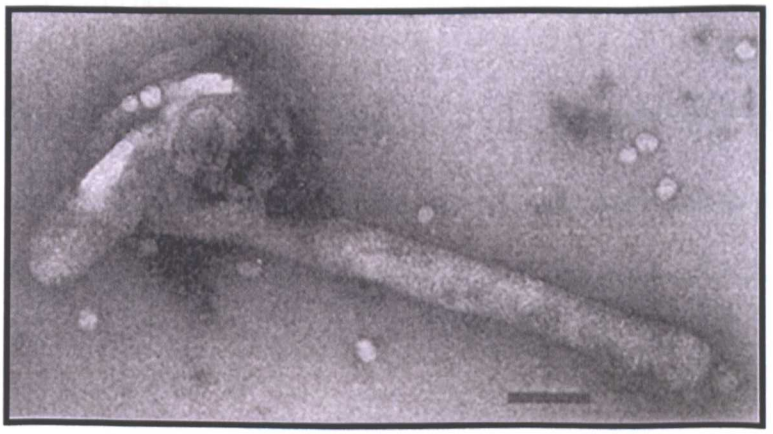

Figure 39: Detection of filovirus-like particles by negative staining electron microscopy. The black bar represents 100 nm.

This suggests further that VP40 assembles at the plasma membrane in a process, which is capable of pinching off vesicles that resemble virus-like structures. Only very few of these filovirus-like particles could be found in the cell culture medium. In an attempt to improve the efficiency of virus particle formation, full-length VP40 was also cloned into pBlueBac 4.5 (Invitrogen) for baculovirus expression in Sf9 insect cells. Although VP40 was indeed expressed at a higher level in insect cells (figure 40), the efficiency of virus-particle release was not improved and similar filovirus-like particles were found in the insect cell culture medium.

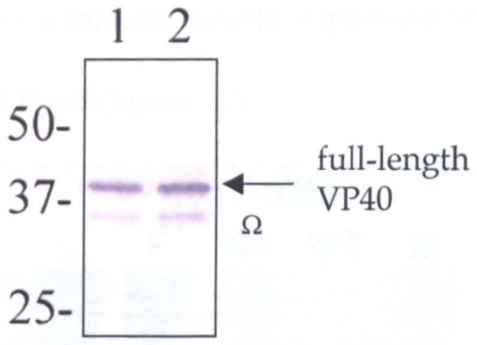

Figure 40: Western blot analysis of VP40 expression in Sf9 insect cells infected with a MOI 1 (lane 1) and 10 (lane 2). The VP40 specific band corresponding to the full-length protein is shown with an arrow, while the degradation product is designated by an asterisk.

Moreover, full-length VP40 appeared to be partially degraded when expressed in insect cells, as opposed to the mammalian cell expressed VP40, which appeared to be more stable.

2. VP40 interacts with two cellular factors

2.1 VP40 and yeast Rsp5p

It has recently become clear that the Proline-rich sequence (PPxY motif) at the N-terminus of Ebola virus VP40 is able to interact with specific WW domains of cellular proteins, including those of yeast Rsp5 protein (Harty et al., 2000). The latter is a member of the large Nedd4 protein family of ubiquitin ligases. Nedd4 orthologues can be found in most eukaryotic organisms. While yeast, mouse and rat proteins have a similar domain organisation human Nedd4 has an additional WW domain that might allow it to interact with other proteins. Preliminary work in the laboratory showed that bacterially expressed VP40 could interact with the WW domain 2 of yeast Rsp5 (figure 41).

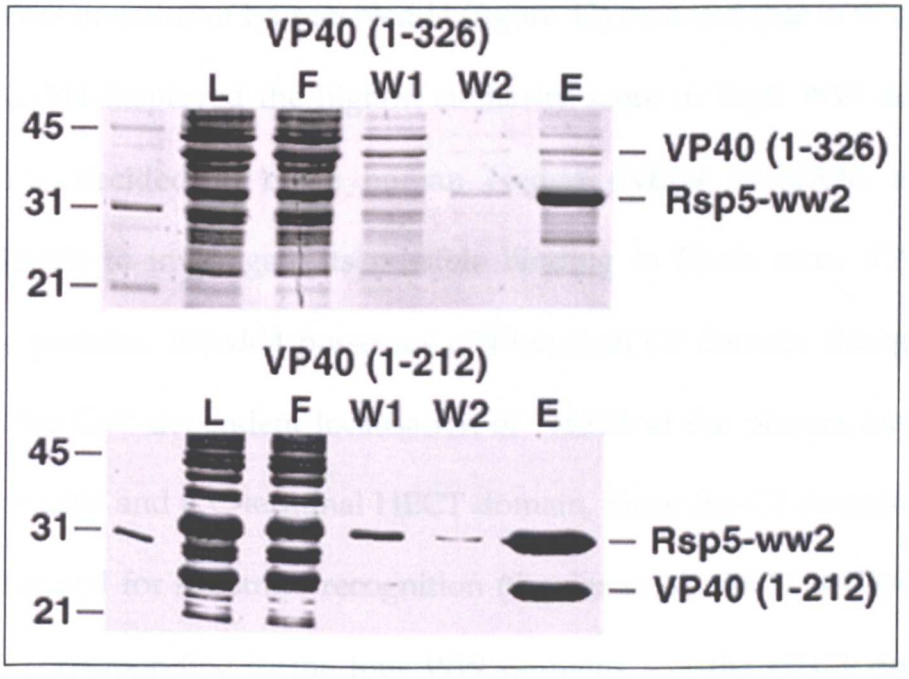

Figure 41: SDS-PAGE analysis of the complex formation between Rsp5 WW domain 2 (as a GST fusion protein) and two forms of Ebola VP40: full-length VP40 (1-326) and a C-terminally truncated form of VP40 (1-212). The cleared supernatant of bacterial extracts expressing VP40 and GST-WW2 (L) were

loaded onto a GST-affinity column and the flow-through (F), washes (W1 and W2) and elution (E) were analysed for the presence of Rsp5 and VP40.

In this experiment, bacteria expressing VP40 and Rsp5 were lysed together and the bacterial extract was loaded on a GST-affinity column. The finding that VP40 co-elutes from the GST-affinity column with the WW domain 2 of Rsp5, in addition to the data presented by Harty *et al.* (Harty et al., 2000) strongly suggests that VP40 can interact with Nedd4-like proteins.

We were therefore interested in investigating the possible interactions between VP40 and a mammalian Nedd4 molecule. A BLAST search of the protein database (SWISSPROT) revealed that human Nedd4 was the closest homologue to yeast Rsp5 WW domain 2 (Figure 42). The alignment of Rsp5 WW domain 2 with each of the WW domains of human Nedd4 (figure 42) revealed that WW domain 3 of human Nedd4 displayed the highest similarity score to Rsp5 WW domain 2. We, therefore, decided to clone human Nedd4 (NED4_HUMAN; accession number: D42055) to investigate its possible binding to Ebola virus VP40. Like other Nedd4 proteins, hNedd4 possesses a N-terminal C2 domain, thought to be involved in the Ca^{2+}-dependent localisation of Nedd4 at the plasma membrane, four WW domains and a C-terminal HECT domain. Since the C2 domain is most likely not required for substrate recognition (Snyder et al., 2001), a cDNA clone (Nedd4ΔC2) corresponding to the four WW domains and the HECT domain of human Nedd4 (2130 base pairs) was amplified as two fragments by polymerase chain reaction from an amplified skeletal muscle cDNA library. The two fragments were subsequently ligated together and subcloned into the pMal vector

(New England Biolabs) for expression as a maltose binding protein (MBP)-fusion protein in *E. coli*.

Results of a BLAST search of SWISSPROT with RSP5 WW2 domain sequence:

Sequences producing High-scoring Segment Pairs:	High Score	Smallest Sum Probability P(N)	N
swiss|P39940|RSP5_YEAST Ubiquitin--protein ligase RSP5 (E...	190	2.5e-14	1
swiss|Q92462|PUB1_SCHPO Ubiquitin--protein ligase pub1 (E..	177	5.6e-13	1
swiss|P46934|NED4_HUMAN NEDD-4 protein (EC 6.3.2.-) (Frag...	133	3.7e-08	1
swiss|P46938|YA65_MOUSE 65 kDa Yes-associated protein (YA...	114	1.6e-06	1
swiss|O95817|BAG3_HUMAN BAG-family molecular chaperone re...	112	3.5e-06	1
swiss|Q9JLV1|BAG3_MOUSE BAG-family molecular chaperone re...	112	3.5e-06	1
swiss|P46936|YA65_CHICK 65 kDa Yes-associated protein (YA...	110	4.0e-06	1
swiss|P46935|NED4_MOUSE NEDD-4 protein (EC 6.3.2.-) (Frag...	110	1.1e-05	1

Alignments of RSP5 WW2 domain with human Nedd4 WW domains:

```
 Score = 133 (51.9 bits), Expect = 3.7e-08, P = 3.7e-08
 Identities = 21/31 (67%), Positives = 24/31 (77%)

Query:         2 LPSGWEQRFTPEGRAYFVDHNTRTTTWVDPR 32
                 LP GWE R  P GR +F+DHNT+TTTW DPR
hNedd4 WW3: 450 LPKGWEVRHAPNGRPFFIDHNTKTTTWEDPR 480

 Score = 115 (45.5 bits), Expect = 3.1e-06, P = 3.1e-06
 Identities = 18/30 (60%), Positives = 23/30 (76%)

Query:         2 LPSGWEQRFTPEGRAYFVDHNTRTTTWVDP 31
                 LP GWE++    GR+Y+VDHN+RTTTW  P
hNedd4 WW2: 377 LPPGWEEKQDERGRSYYVDHNSRTTTWTKP 406

 Score = 101 (40.6 bits), Expect = 9.5e-05, P = 9.5e-05
 Identities = 15/31 (48%), Positives = 22/31 (70%)

Query:         2 LPSGWEQRFTPEGRAYFVDHNTRTTTWVDPR 32
                 LP GWE+R   +GR ++++HN + T W DPR
hNedd4 WW4: 502 LPPGWEERTHTDGRIFYINHNIKRTQWEDPR 532

 Score = 86 (35.3 bits), Expect = 0.0038, P = 0.0038
 Identities = 15/30 (50%), Positives = 19/30 (63%)

Query:         2 LPSGWEQRFTPEGRAYFVDHNTRTTTWVDP 31
                 LP GWE+R    GR Y+V+H +R T W  P
hNedd4 WW1: 220 LPPGWEERQDILGRTYYVNHESRRTQWKRP 249
```

Figure 42: Results of the BLAST search of the SWISSPROT database with Rsp5 WW domain 2, and sequence alignment of Rsp5 WW domain 2 with each of the human Nedd4 WW domains.

2.2 VP40 and human Nedd4

In the crystallised form of VP40 (missing the first 30 residues), the protein is clearly monomeric in solution. However, upon trypsin digestion of VP40 (31-326) at residue 212, the protein oligomerises and forms ring-like structures (Ruigrok et al., 2000b). Thus, removal of most of the C-terminal domain is able to provoke the conformational change necessary for oligomerisation. In fact, it is now clear that removal of the last 7 residues of VP40 (1-319) is sufficient to destabilise the two domains, and allow oligomerisation to occur (Scianimanico et al., 2000).

In this study we have, therefore, purified several different forms of VP40 with or without the full N-terminus and with different C-terminal truncations (figure 43). All the VP40 constructs were expressed in BL21 pUBS and expression was induced by addition of 1 mM IPTG.

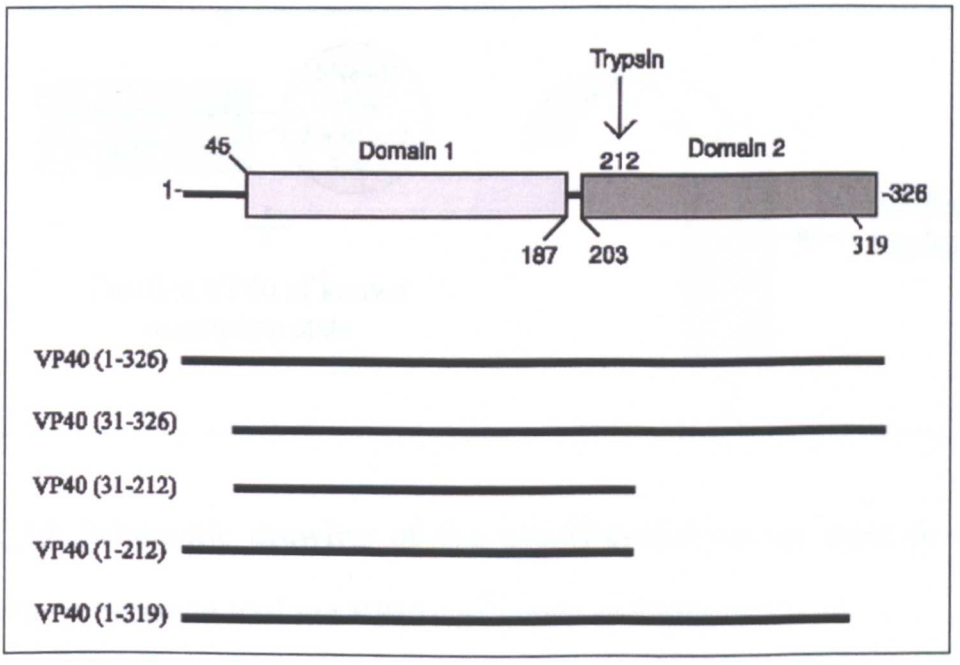

Figure 43: Summary of the VP40 constructs used for the binding assay with hNedd4. These VP40 cDNAs were cloned into pRSET (Invitrogen) and expressed in *E. coli*.

Several chromatography steps were needed to obtain pure VP40 protein (as described in Materials and Methods). The purified proteins were then subjected to a size exclusion column (Superdex 200; Pharmacia) so as to determine their oligomeric state. The superdex 200 was calibrated with known molecular weight markers: ferritin (440 kDa) elutes at 10 ml, while ovalbumin (43 kDa) elutes at 14 ml. Monomeric VP40 elutes at 14-15 ml, while oligomeric VP40 elutes earlier at 10-11 ml.

In parallel, WW domains 2, 3 and 4 of hNedd4 were expressed as an MBP fusion protein (Nedd4 (WW234)-MBP) and purified on amylose resin (Figure 45, lane 1). WW domain 1 of hNedd4 was not used in this work because it appeared to be very insoluble in all buffer conditions tested. To test the binding of the various VP40 constructs to hNedd4 WW domains, the experimental set-up shown in figure 44 was followed.

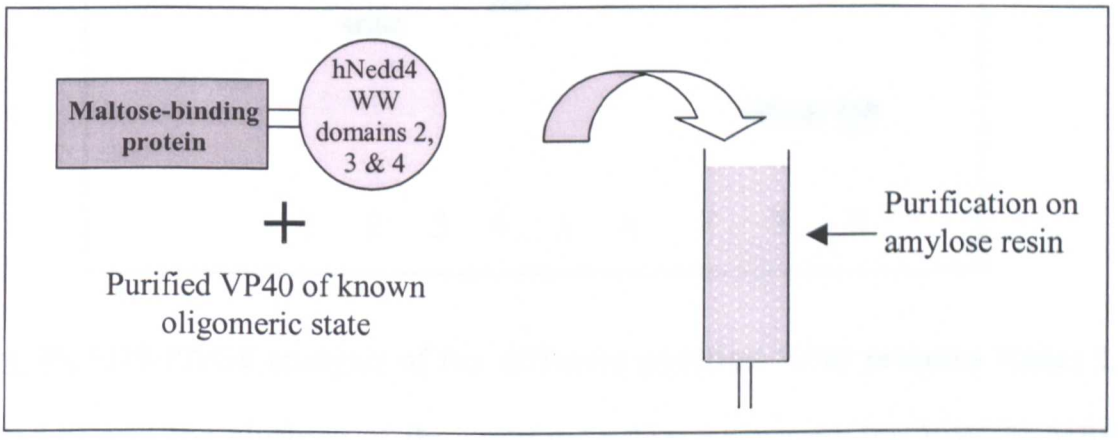

Figure 44: Schematic drawing of the experimental set-up used to study the interactions between various VP40 constructs and hNedd4.

In turn, each purified VP40 (Figure 45, lanes 2, 4, 6 and 8) was mixed with Nedd4 (WW234)-MBP and loaded on the amylose resin as shown in figure 44. The elutions from these binding assays are shown in figure 45, lanes 3, 5, 7 and 9. In

figure 45, lane 1 corresponds to the purified WW234-MBP to which each of the VP40 proteins was added. These results (summarised in Table 5) clearly show that a full N-terminus of VP40 including the PPxY motif is required for binding to hNedd4 WW domains, since there is no coelution of Nedd4 (WW234)-MBP with VP40 (31-326). On the other hand, we observe that the C-terminally truncated forms of VP40 (VP40 (1-212) and VP40 (1-319)) clearly interact with Nedd4 (WW234)-MBP.

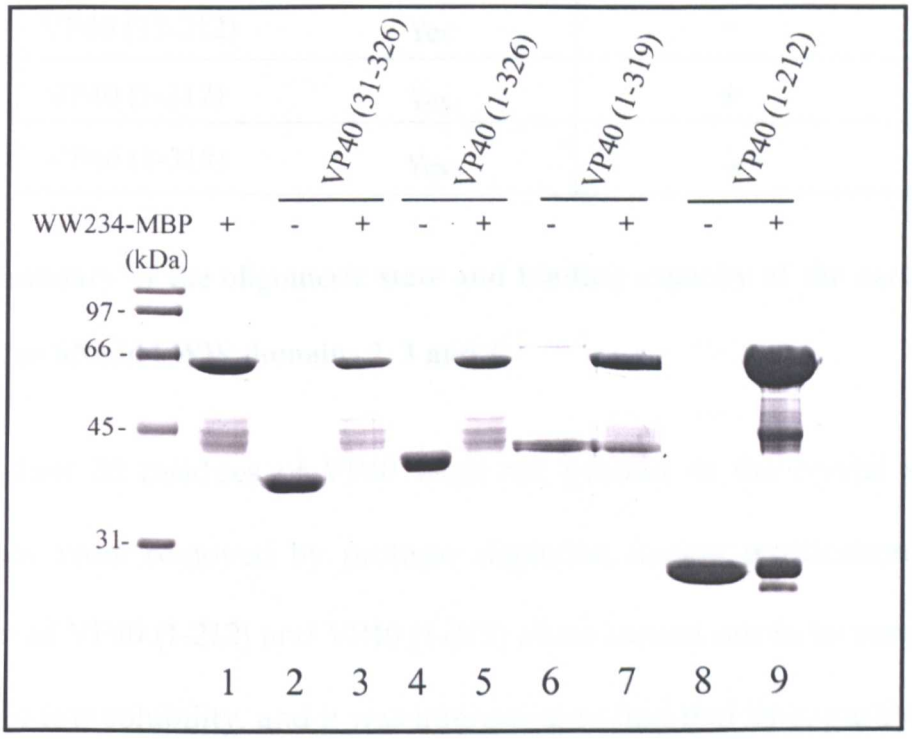

Figure 45: SDS-PAGE analysis of the different purified VP40 proteins (lanes 2, 4, 6 and 8) and the elutions of the amylose column after mixing WW234-MBP (lane 1) with each of the VP40 constructs (lanes 3, 5, 7 and 9).

These C-terminally truncated forms of VP40, unlike the full-length VP40, are oligomeric in solution as can be seen from the gel filtration elution profiles (data not shown). Interestingly, from figure 45, we observe that full-length VP40 (containing the N-terminal PPxY motif), which is a monomer in solution, does not coelute with Nedd4 (WW234)-MBP. These results thus suggest that, *in vitro*, the

ring-like form of VP40, which has been suggested to constitute the building blocks for virus assembly *in vivo*, has a higher affinity for human Nedd4 WW domains, than does monomeric VP40.

	Forms oligomers	Binds to WW domains 2, 3 & 4 of hNedd4
VP40 (1-326)	No	
VP40 (31-326)	No	-
VP40 (31-212)	Yes	-
VP40 (1-212)	Yes	+
VP40 (1-319)	Yes	+

Table 5: Summary of the oligomeric state and binding capacity of the various VP40 constructs for hNedd4 WW domains 2, 3 and 4.

The first 30 residues of VP40 were not present in the crystal structure, because they were removed by protease digestion during purification. Indeed, purification of VP40 (1-212) and VP40 (1-319) alone turned out to be very difficult as a result of low solubility, and it was interesting to find that by copurifying these two VP40 constructs with Nedd4 (WW234)-MBP, the N-terminus of VP40 was stabilised. It thus appears that the binding of human Nedd4 WW domains to the PPxY motif of VP40 improves the proper folding of the protein *in vitro*.

We, then, prepared a number of Nedd4 deletion constructs (shown in figure 46) for expression as MBP fusion proteins. We progressively removed domains, until we abolished the binding of Nedd4 to VP40. The VP40 construct used for this study was VP40 (1-212), which contains a full N-terminal domain and is

oligomeric in solution. It thus displayed a higher binding affinity for WW domains 2, 3 and 4 (figure 45, lane 9).

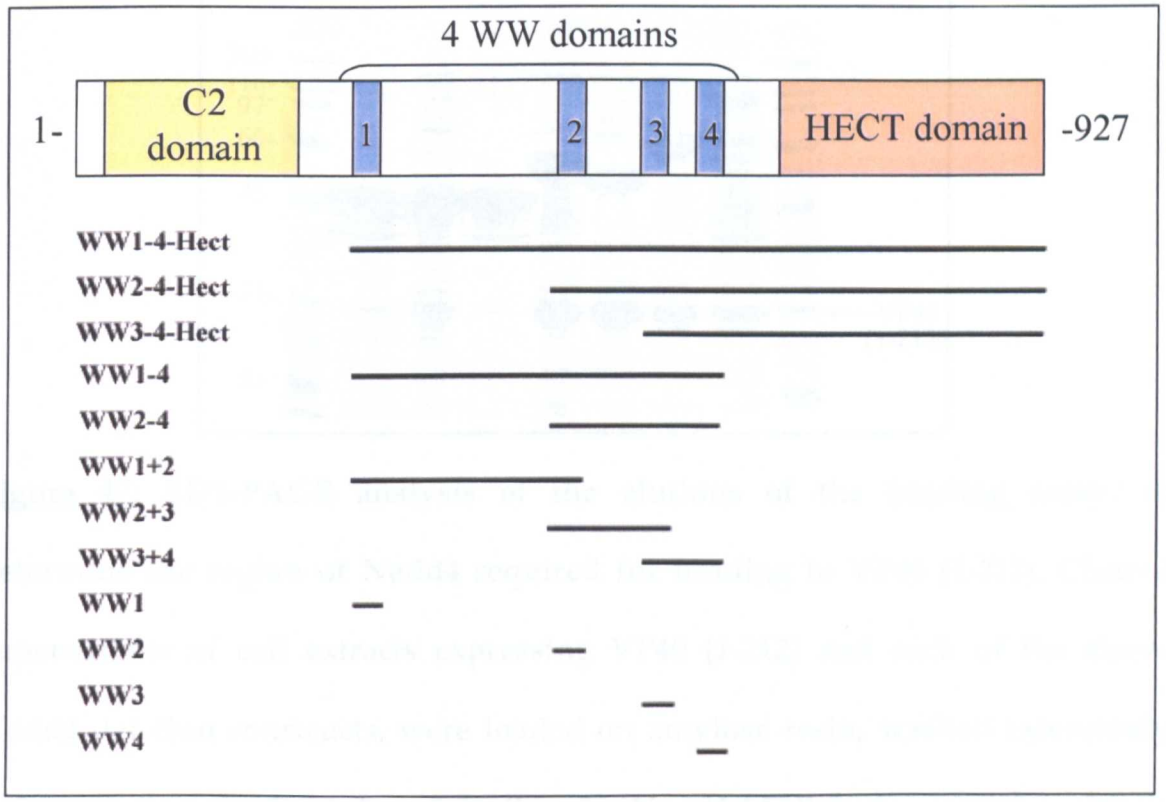

Figure 46: Human Nedd4 domain organisation. The deletion constructs of hNedd4 used for expression are shown below.

The results of this binding assay are shown in Figure 47. It appeared very clearly from these results that the presence of WW domain 3 was a prerequisite for binding of Nedd4 to VP40 *in vitro*. Indeed, WW domain 3 alone, unlike WW domains 2 and 4 alone, was sufficient for efficient binding to VP40 (1-212). In the cases of WW domains 2 and 4, we observed some very faint binding, but there was a clear difference in affinity of each WW domain for VP40. WW domain 3 is, thus necessary and sufficient for binding of hNedd4 to VP40. It has already been observed that different WW domains from a given Nedd4 protein possess differential substrate specificity *in vitro* (Sudol, 1996).

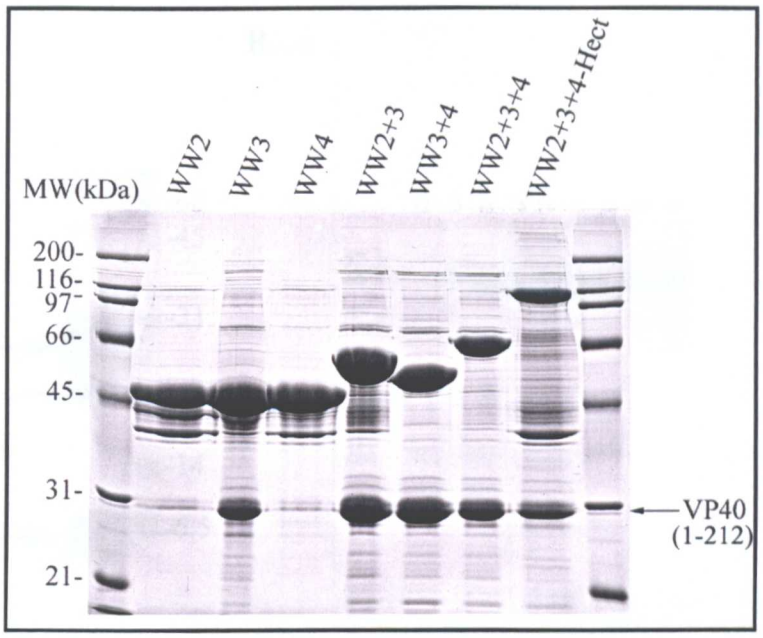

Figure 47: SDS-PAGE analysis of the elutions of the binding assays to determine the region of Nedd4 required for binding to VP40 (1-212). Cleared supernatants of cell extracts expressing VP40 (1-212) and each of the above Nedd4 deletion constructs, were loaded on amylose resin, washed extensively before eluting the bound protein (i.e. the Nedd4-MBP fusion protein with or without VP40 (1-212)).

In view of these results, a number of crystallisation trials were undertaken. In order to better understand the binding specificity of VP40 (1-212) for hNedd4 WW domain 3, crystallisation trials were set up with purified VP40 (1-212) and WW domain 3, with an excess of the WW domain. The two proteins were mixed (figure 48A) and concentrated together to 8 mg/ml prior to setting up the crystallisation drops. In addition, a construct of hNedd4 corresponding to the three WW domains 2, 3 and 4 and the HECT domain was also used for some initial crystallisation trials (figure 48B), since there is currently no known structure of multiple domains of a Nedd4 protein. So far no crystals were obtained.

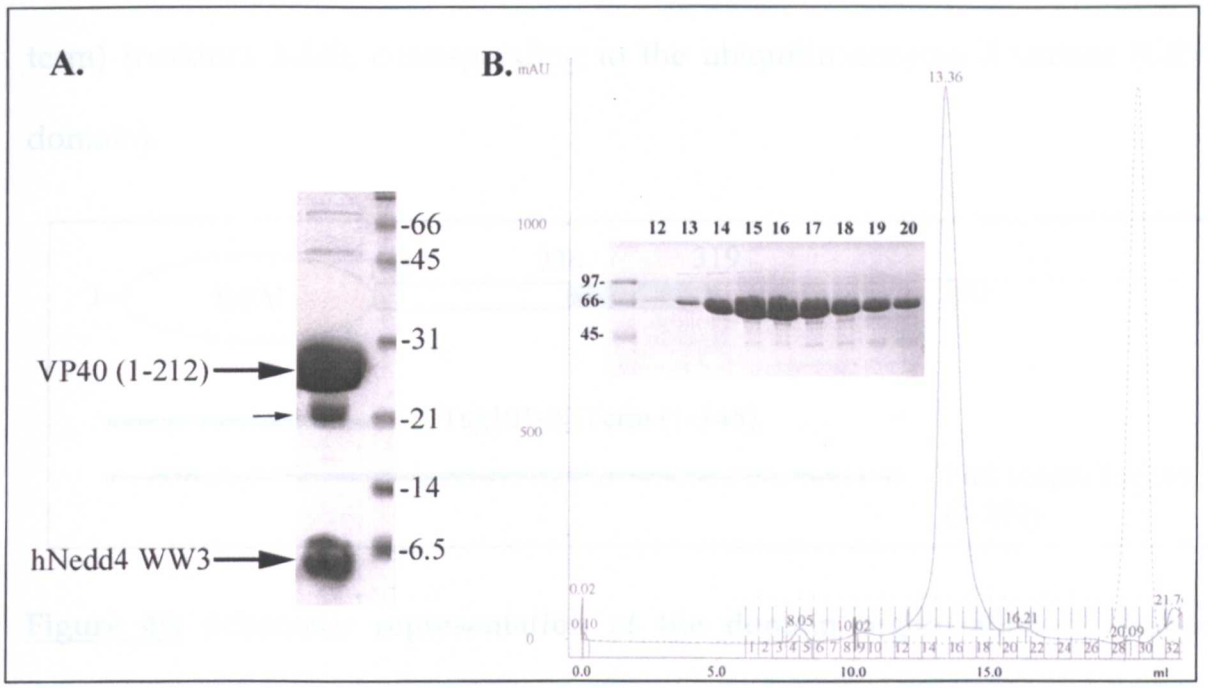

Figure 48: (A) Concentrated VP40 (1-212) and human Nedd4 WW domain 3 used for crystallisation trials. The small arrow indicates most probably a degradation product of VP40 (1-212). (B) Size exclusion chromatography of human Nedd4 (WW domains 2, 3 and 4 and the HECT domain) on a Superdex 200. The Nedd4 protein elutes at 13.4 ml and the SDS-PAGE analysis of the protein fractions corresponding to the peak is also shown.

2.3 VP40 and human Tsg101

Over the past year, several studies have led to the identification of a new cellular factor required for efficient viral budding, namely Tsg101 (Garrus et al., 2001; Luban, 2001; Martin-Serrano et al., 2001; Perez and Nolan, 2001). In HIV-1, the PTAP sequence within p6-Gag protein was shown to be required for binding to Tsg101. Interestingly, VP40 also possesses this motif at its N-terminus (chapter 4, figure 30). We therefore cloned the full-length human Tsg101 (1170 base pairs; accession number: U82130) from a skeletal muscle cDNA library and two expression constructs (figure 49) were prepared, as MBP-fusion proteins: full-length Tsg101 (residues 1-390) and a C-terminally truncated form, Tsg101 (N-

term) (residues 1-145, corresponding to the ubiquitin enzyme 2 variant (UEV) domain).

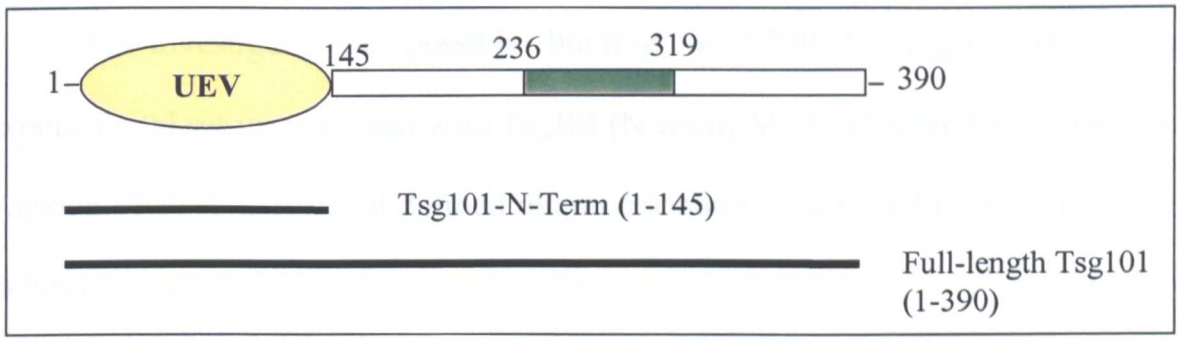

Figure 49: Schematic representation of the domain organisation of human Tsg101. The yellow N-terminal domain corresponds to a ubiquitin enzyme 2 variant (UEV) domain, and the green rectangle represents a region predicted (using the SMART programme) to form a coiled coil. The constructs used for bacterial expression are shown as black bars below.

Only the truncated form was detectably expressed in bacteria and was therefore used for the further binding studies (figure 50).

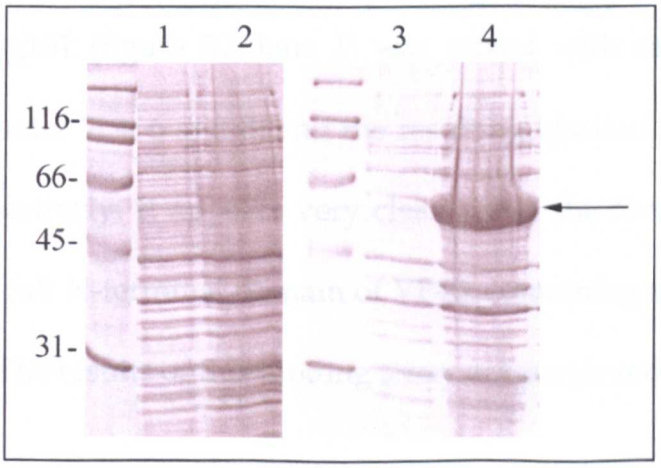

Figure 50: SDS-PAGE analysis of Tsg101 expression in *E. coli*. Lanes 1 and 2 correspond to the full-length clone of Tsg101 before and after induction with 1 mM IPTG, and lanes 3 and 4 correspond to the C-terminally truncated form of

Tsg101 before and after induction with 1 mM IPTG. The arrow shows the expression band corresponding to the N-terminal Tsg101 fused to MBP.

To investigate the possible binding of VP40 to Tsg101, the same experimental set-up was used with Tsg101 (N-term)-MBP as for Nedd4 (figure 51). Various VP40 constructs, of defined oligomeric state were tested for their capacity to bind to Tsg101.

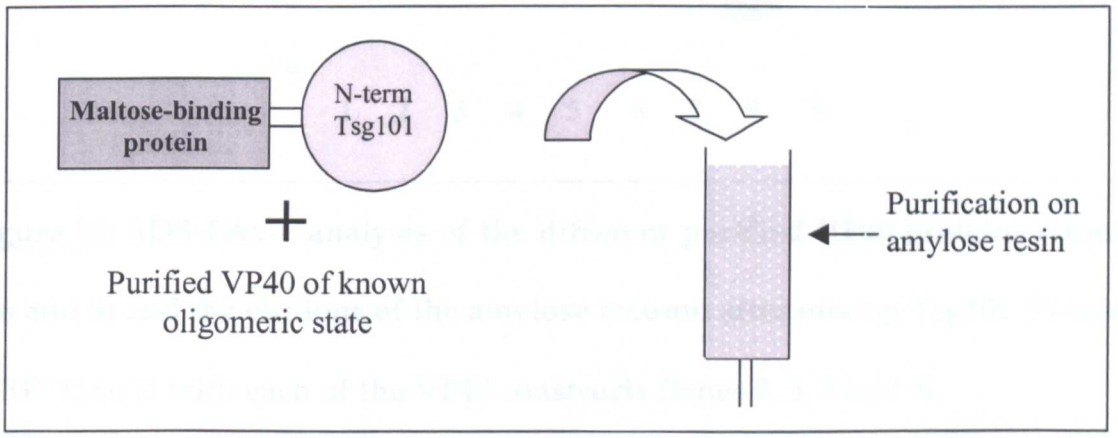

Figure 51: Schematic drawing of the experimental set-up used to study the interactions between various VP40 constructs and hTsg101.

Purified Tsg101 (figure 52, lane 1) was mixed with each of the purified VP40 constructs (lanes 2, 4, 6 and 8) and the resulting elutions are shown in lanes 3, 5, 7 and 9 respectively. It appears very clearly that the N-terminal domain of Tsg101 requires a full N-terminal domain of VP40, containing the PTAP motif, for efficient binding. The results of this binding assay are presented in figure 53.

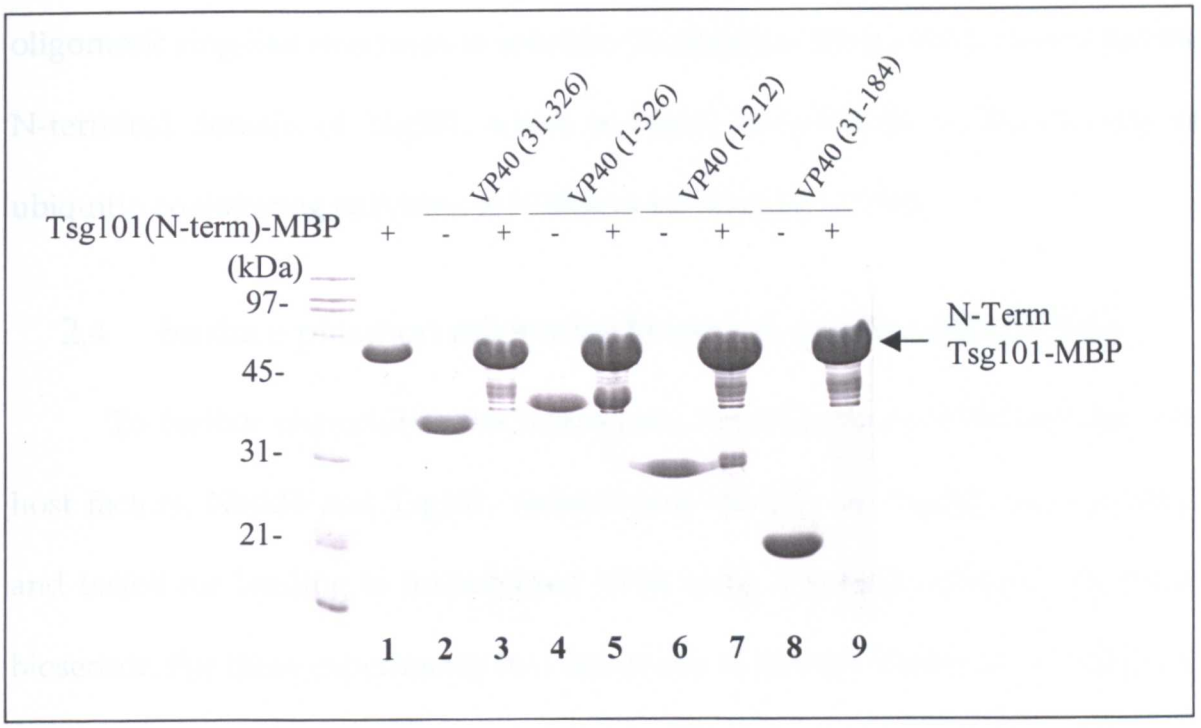

Figure 52: SDS-PAGE analysis of the different purified VP40 proteins (lanes 2, 4, 6 and 8) and the elutions of the amylose column after mixing Tsg101 (N-term)-MBP (lane 1) with each of the VP40 constructs (lanes 3, 5, 7 and 9).

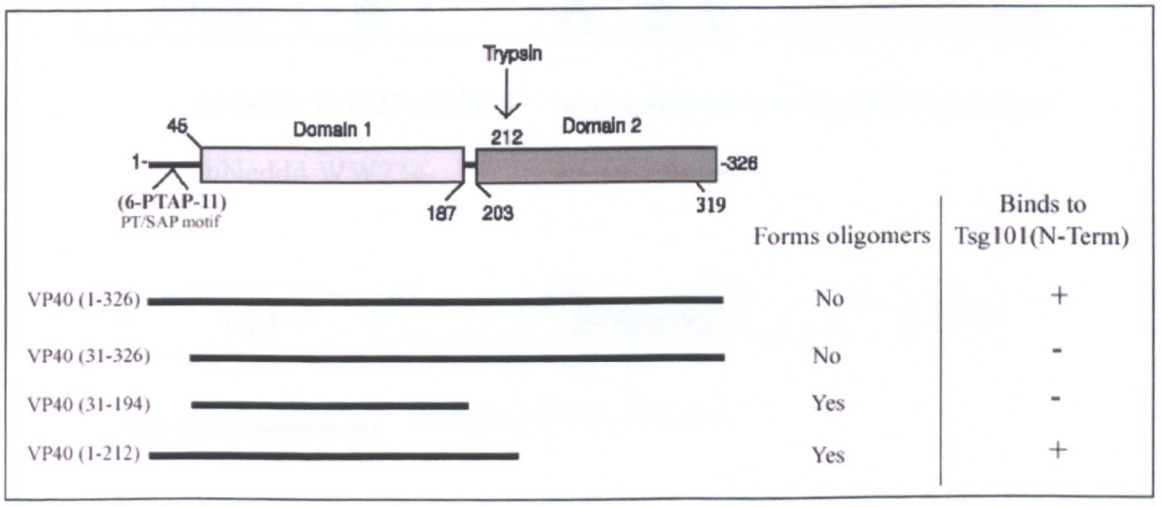

Figure 53: Summary of the binding experiments of various VP40 constructs, of defined oligomeric state, with the N-terminal domain of Tsg101.

Unlike Nedd4, human Tsg101 however appears to bind equally well to monomeric and oligomeric VP40, since Tsg101 coelutes with both full-length VP40 (1-326), which is a monomer in solution, and with VP40 (1-212), which forms

oligomeric ring-like structures in solution. In addition, these results reveal that the N-terminal domain of Tsg101, which is highly homologous to the E2 class of ubiquitin conjugating enzymes, is sufficient for binding to VP40.

2.4 Surface plasmon resonance biosensor studies (BIACORE)

To further characterise the interactions between Ebola VP40 and the two host factors, Nedd4 and Tsg101, recombinant Nedd4 and Tsg101 were purified and tested for binding to immobilised VP40 using a surface plasmon resonance biosensor. For these experiments two constructs of human Nedd4 were tested, the UEV domain of Tsg101 and two forms of VP40, cloned into pET21d with a C-terminal His-tag (figure 54).

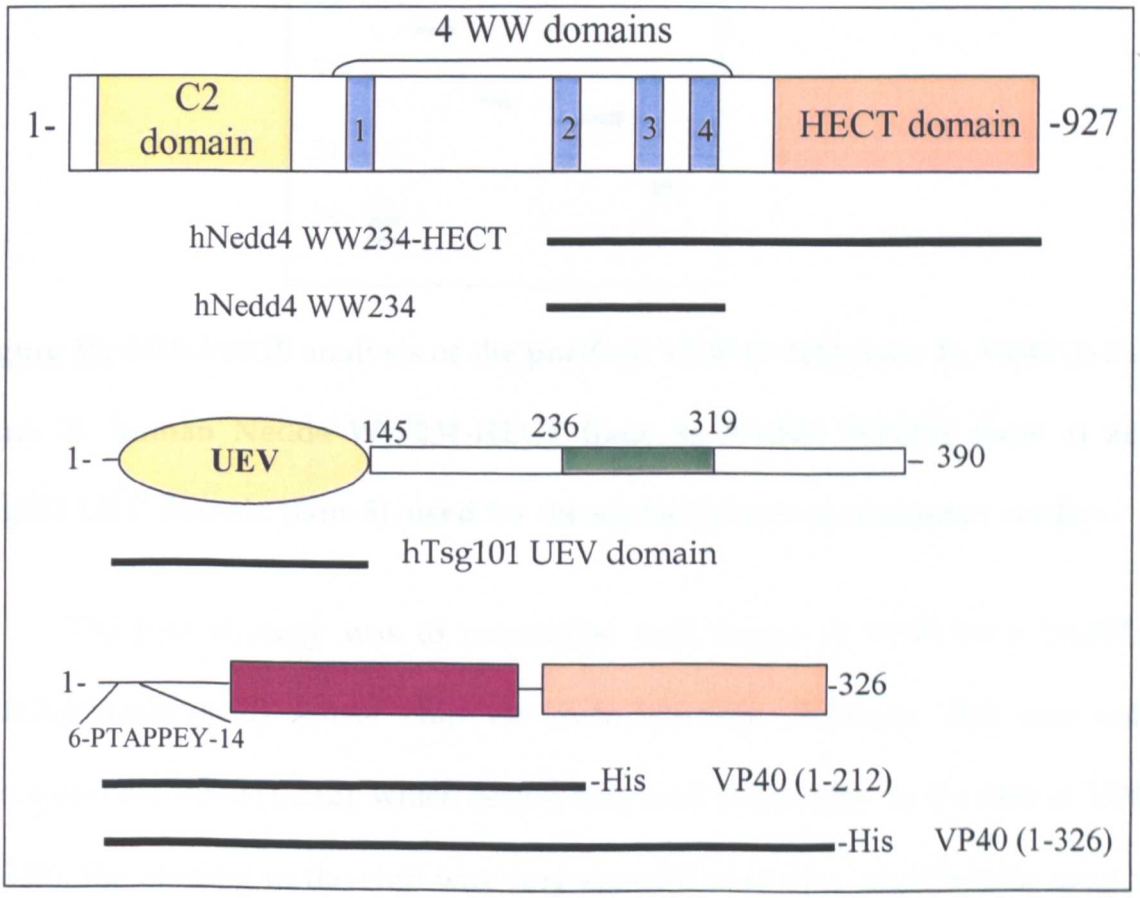

Figure 54: Schematic diagram of the various constructs of Nedd4, Tsg101 and VP40 used for the surface plasmon resonance studies.

For these studies, each of the proteins was purified to homogeneity (figure 55). The Nedd4 and Tsg101 constructs were expressed as MBP-fusion proteins and the MBP moiety was removed from both Nedd4 and Tsg101 by cleavage with the TEV protease. The VP40 constructs were prepared with a C-terminal His-tag, so that the tag would not interfere with the possible binding of either Nedd4 or Tsg101 to the protein-protein binding motifs found at the N-terminus of VP40. To ensure that the His-tag was not removed by proteolytic digestion during the purification steps, the His-tagged protein was analysed by Western blot using an anti-His monoclonal antibody.

Figure 55: SDS-PAGE analysis of the purified VP40 (1-326) (lane 1), VP40 (1-212) (lane 2), human Nedd4 WW234-HECT (lane 3), Nedd4 WW234 (lane 4) and Tsg101 UEV domain (lane 5), used for the surface plasmon resonance studies.

The first strategy was to immobilise both forms of VP40 on a Ni-NTA (nitrilotriacetic acid) sensor chip via their His-tags. However, this was only successful for VP40 (1-212), which bound very well to the chip. In the case of VP40 (1-326), the binding to the chip was very unstable and thus could not be used to analyse the binding of Nedd4 and Tsg101. A stable baseline is required to obtain a clear binding signal. Nevertheless, this experimental set-up was used to study the binding of VP40 (1-212) with Nedd4 WW234 and WW234-HECT, and with Tsg101

UEV domain. In all cases, there was clear evidence of binding of each of these proteins to VP40 (1-212). The sensorgrams corresponding to the binding of Nedd4 WW234 and WW234-HECT to VP40 (1-212) are shown in figure 56. In these experiments, 200 nM of VP40 (1-212) were immobilised on the Ni-NTA chip and the Nedd4 proteins were injected at a concentration of 150 μg/ml. The binding pattern is similar in both cases, implying that the HECT domain does not significantly influence the complex formation, although there does seem to be slightly more Nedd4 WW234-HECT bound to VP40 (1-212), compared to Nedd4 WW234. This difference may result from the higher molecular weight of Nedd4 WW234-HECT, relative to WW234. In both cases the complex formation and dissociation are very rapid.

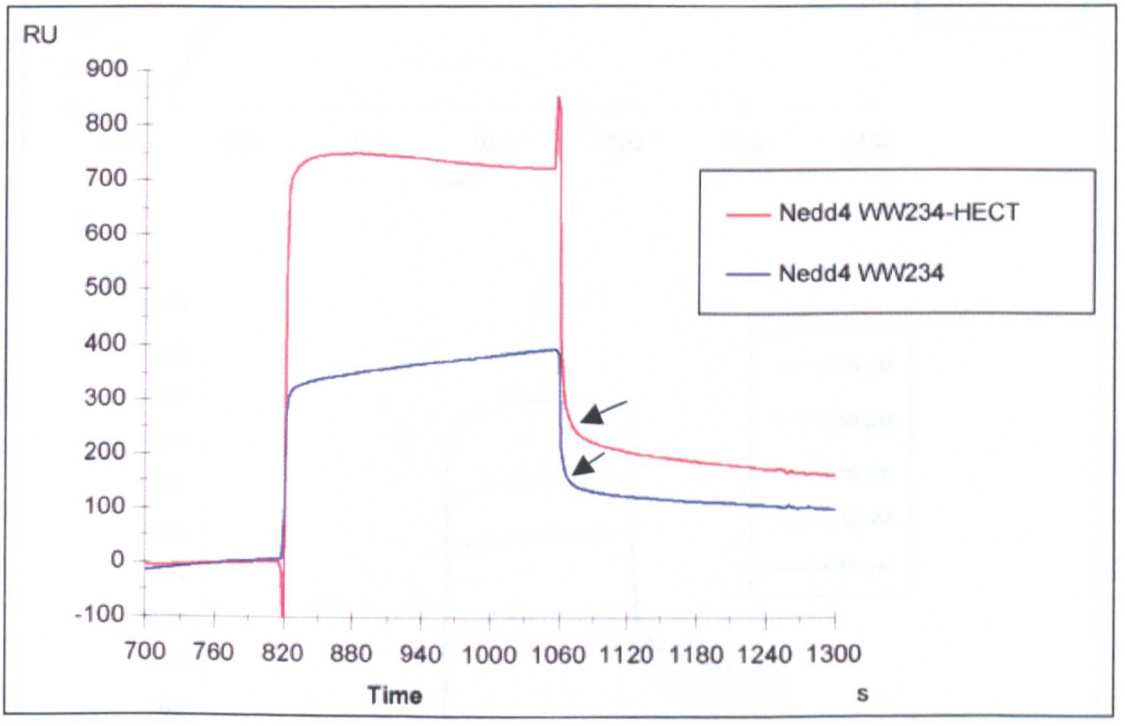

Figure 56: Sensorgram obtained after injection of either Nedd4 WW234 (blue) or Nedd4 WW234-HECT (red) on immobilised VP40 (1-212). RU: response units, defined as 1000 RU=1 ng/mm². The arrows point to the area of the curve used to determine the amount of Nedd4 bound to VP40 (1-212) after washing away the excess Nedd4.

Different concentrations of Nedd4 WW234-HECT were injected on immobilised VP40 (1-212) to estimate the binding affinity (figure 57A and B). In between each consecutive injection of Nedd4, the sensor chip was regenerated and more VP40 (1-212) was immobilised on the surface.

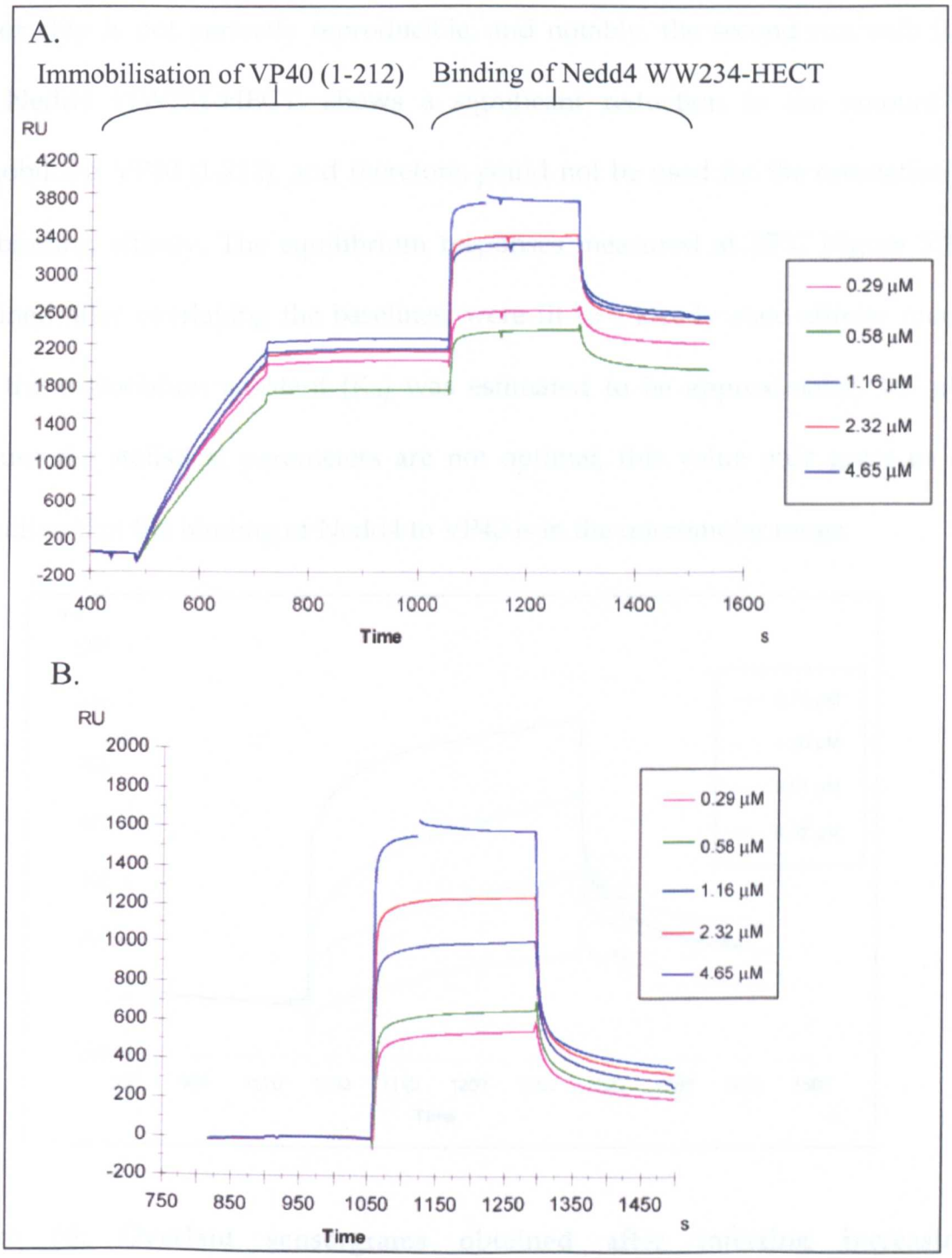

Figure 57: Sensorgrams obtained after the injection of increasing concentrations (0.29, 0.58, 1.16, 2.32 and 4.65 µM) of Nedd4 WW234-HECT on immobilised VP40

(1-212). (A) Crude sensorgrams, showing the immobilisation of VP40 (1-212) followed by the injection of the ligand, Nedd4 WW234-HECT. (B) Overlaid sensorgrams used to estimate the binding affinity.

As we can see from figure 57A, the amount of VP40 (1-212) bound to the sensor chip is not perfectly reproducible, and notably, the second run with 0.58 μM Nedd4 WW234-HECT, shows a significant reduction in the amount of immobilised VP40 (1-212), and therefore, could not be used for the estimation of the binding affinity. The equilibrium responses measured at 25°C (figure 57B), obtained after overlaying the baselines, were fit to a steady state affinity model and the dissociation constant (K_d) was estimated to be approximately 0.8 μM. Because the statistical parameters are not optimal, this value only gives us an indication that the binding of Nedd4 to VP40 is in the micromolar range.

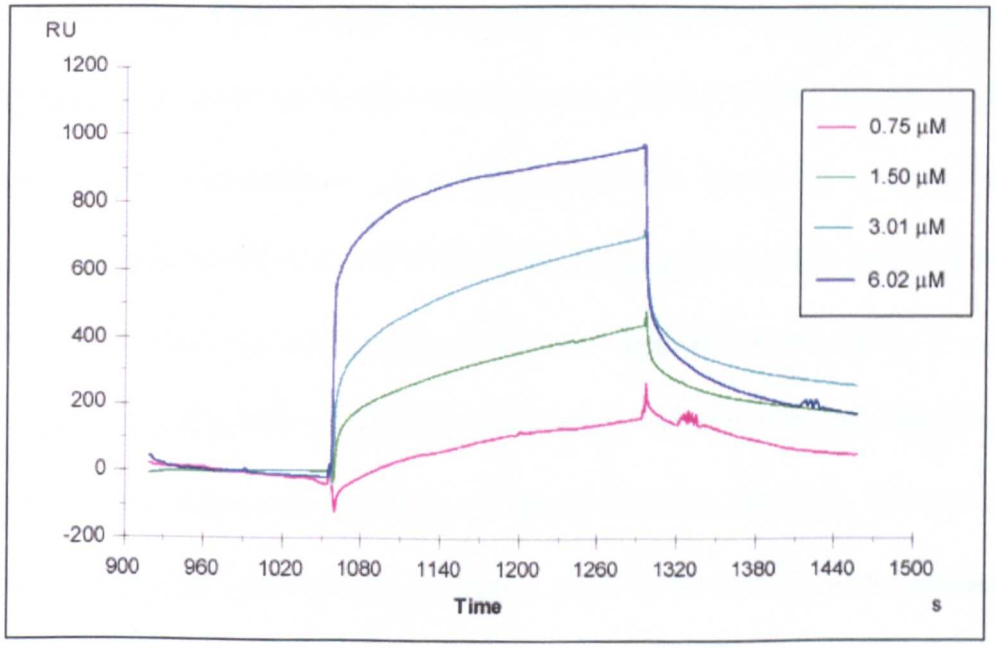

Figure 58: **Overlaid sensorgrams obtained after injecting increasing concentrations of the UEV domain of Tsg101 (0.75, 1.50, 3.01 and 6.02 μM) on immobilised VP40 (1-212).**

Different concentrations (0.75, 1.50, 3.01 and 6.02 μM) of the UEV domain of Tsg101 (figure 58) were also injected on immobilised VP40 (1-212), which clearly showed that Tsg101 could specifically bind to VP40 (1-212), but no reliable affinity constant could be determined from this data. The sensorgrams shown in figure 58 could not be fit to any of the following reaction models using the BIAevaluation™ software: Langmuirian 1:1 complex formation, heterogeneous ligand, conformational change and heterogeneous analyte. The complex between Tsg101 and VP40 (1-212) appears to form very rapidly (as in the case of Nedd4), while the dissociation appears to be slower than in the case of Nedd4.

In order to compare the possible interactions of VP40 (1-326) and VP40 (1-212) with Nedd4 and Tsg101, both VP40 proteins were dialysed against a maleate buffer (pH 6) so as to immobilise the proteins covalently via their $-NH_2$ groups on a CM5 sensor chip. VP40 (1-212) largely precipitated during the dialysis step, and was therefore not used for further experiments. VP40 (1-326), on the other hand, was successfully immobilised on a CM5 chip and could be used to study its binding to Nedd4 WW234 and WW234-HECT (figure 59A). The binding of Nedd4 to VP40 (1-326) was very poor, although it was slightly improved by the presence of the HECT domain (shift in the final amount of Nedd4 WW234-HECT bound in figure 59A). To compare the binding of Nedd4 WW234-HECT to VP40 (1-212) and to VP40 (1-326), the sensorgrams obtained with VP40 (1-212) immobilised via its His-tag (red curve in figure 57A) and with VP40 (1-326) immobilised covalently via its amine groups (figure 59A) were overlaid (figure 59B). VP40 (1-212) clearly binds more Nedd4 WW234-HECT, than does VP40 (1-326).

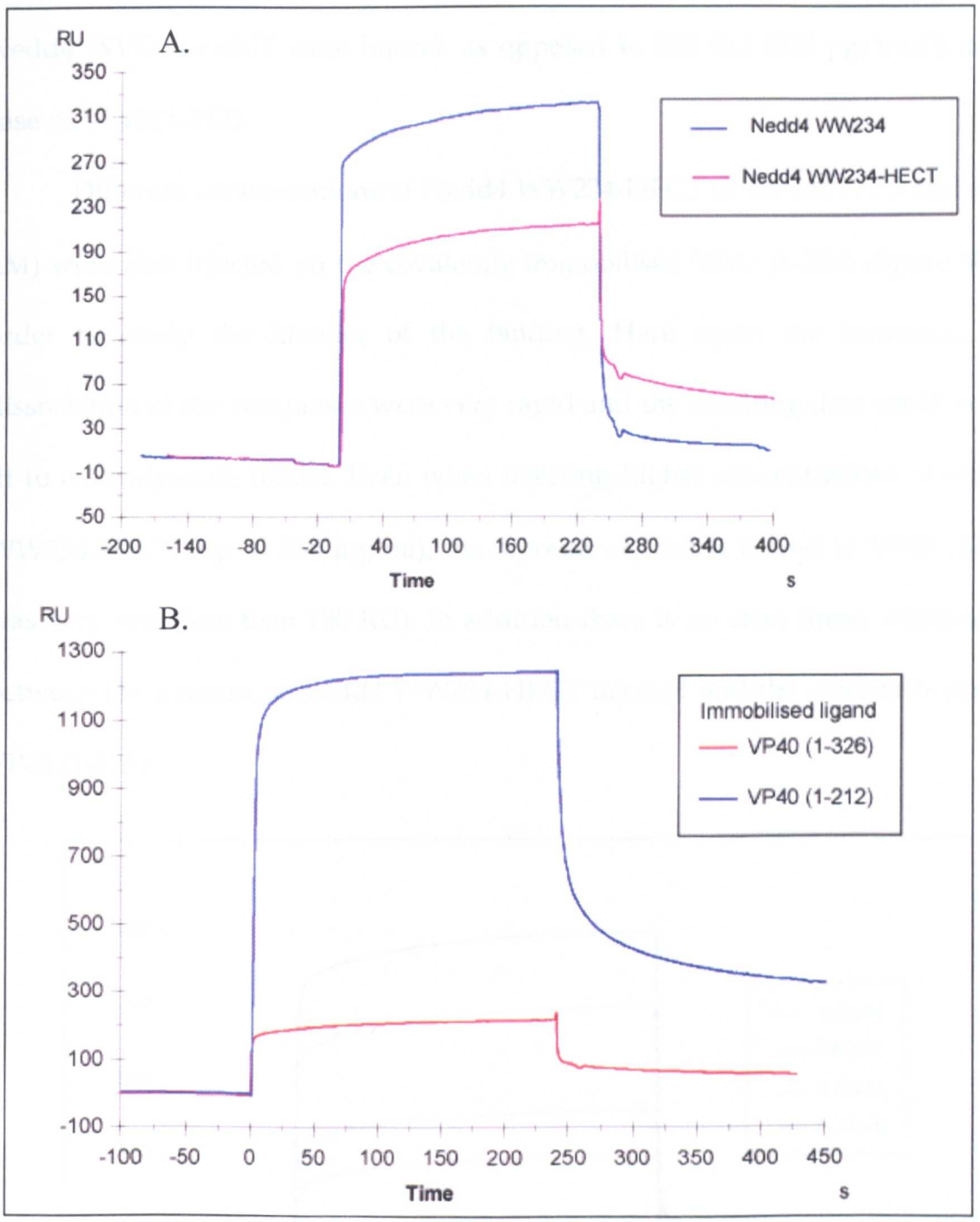

Figure 59: (A) Sensorgram obtained after injection of 150 µg/ml of either Nedd4 WW234 (blue) or Nedd4 WW234-HECT (pink) on covalently immobilised VP40 (1-326). (B) Overlaid sensorgrams obtained after injecting 150 µg/ml of Nedd4 WW234-HECT on either immobilised VP40 (1-212) (blue) or immobilised VP40 (1-326) (red).

Very similar amounts of each of the VP40 molecules were immobilised on each of the sensor chips (91.3 fmoles/mm^2 for VP40 (1-212) and 75.6 fmoles/mm^2

for VP40 (1-326)), and yet in the case of VP40 (1-326) only 93 RU (93 pg/mm^2) of Nedd4 WW234-HECT were bound, as opposed to 520 RU (520 pg/mm^2) in the case of VP40 (1-212).

Different concentrations of Nedd4 WW234-HECT (2.32, 3.87, 7.75 and 11.62 μM) were also injected on the covalently immobilised VP40 (1-326) (figure 60) in order to study the kinetics of the binding. Here again the formation and dissociation of the complexes were very rapid and the resulting data could not be fit to a steady-state model. Even when injecting higher concentrations of Nedd4 WW234-HECT (up to 750 μg/ml), the amount of Nedd4 bound to VP40 (1-326) was very low (less than 150 RU). In addition there is no clear linear relationship between the amount of Nedd4 WW234-HECT injected and the amount bound to VP40 (1-326).

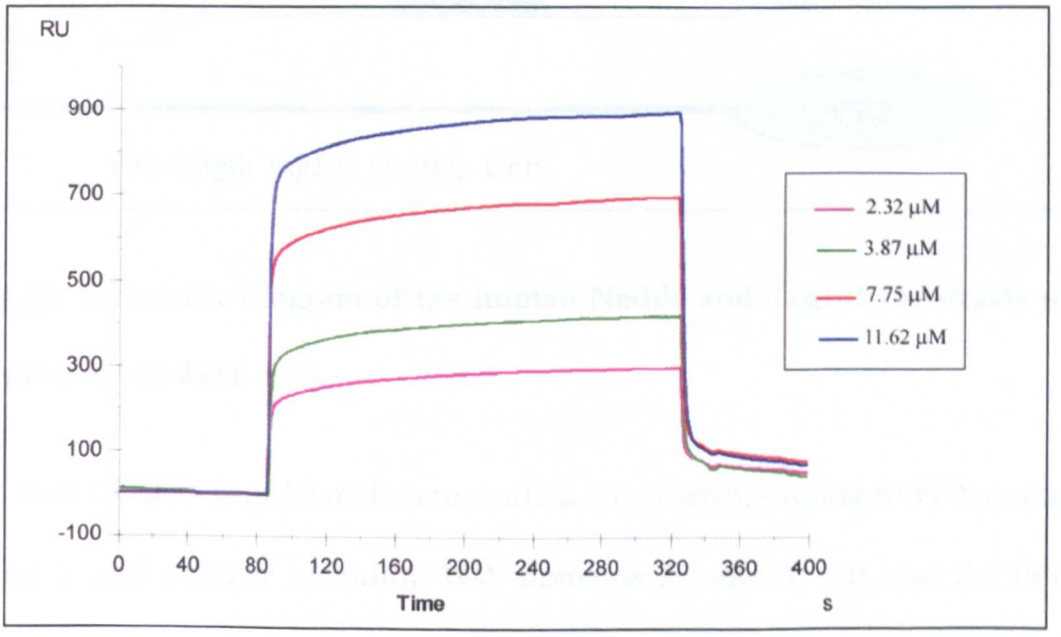

Figure 60: Overlaid sensorgrams obtained after injecting increasing concentrations of Nedd4 WW234-HECT (2.32, 3.87, 7.75 and 11.62 μM) on immobilised VP40 (1-326).

2.5 Cellular localisation of human Nedd4 and Tsg101

In order to further understand the role of these two cellular factors in filovirus budding, it was of interest to look at the cellular localisation of human Nedd4 and Tsg101 in 293T cells. For this, both Nedd4 and Tsg101 were cloned into pcDNA 3.1 for expression in mammalian cells, in this case human embryonic kidney 293T cells.

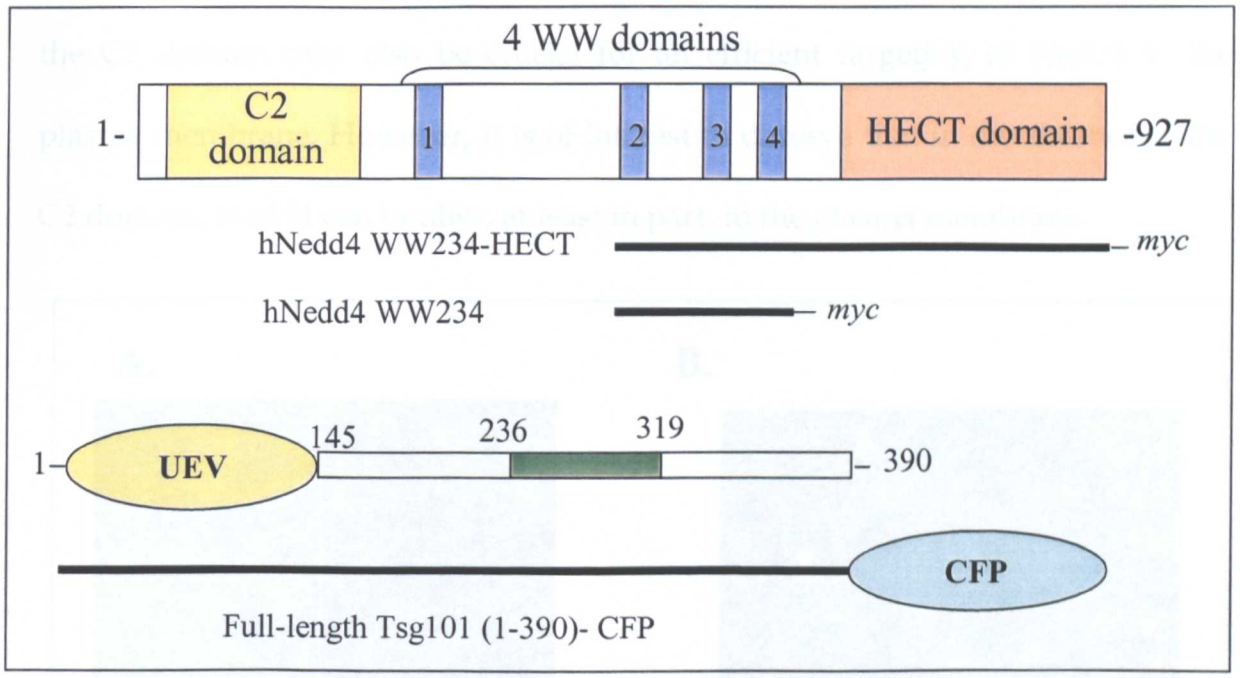

Figure 61: Schematic diagram of the human Nedd4 and Tsg101 constructs used for expression in 293T cells.

Two constructs of Nedd4 were studied, one corresponding to WW domains 2, 3 and 4, and another including WW domains 2,3 and 4, but also the full C-terminal HECT domain (figure 61). The C2 domain was not included in the constructs, since the full DNA sequence corresponding to the C2 domain was not available in the database, and could thus not be amplified from the cDNA library. These two constructs were cloned with an additional C-terminal *myc*-epitope for detection by Western blot and indirect immunofluorescence using a mouse

monoclonal anti-*myc* antibody (figure 62). A full-length clone of human Tsg101 was prepared as a fusion protein with cyan-fluorescent protein (CFP) (figure 61).

Interestingly, we found that the WW domains 2, 3 and 4 of hNedd4 show intense staining of the cytoplasm of transfected cell (figure 62A), while the presence of the HECT domain was sufficient for targeting a small proportion of the protein to the plasma membrane (figure 62B). The C2 domain is believed to be responsible for the regulated presence of Nedd4 at the plasma membrane. Thus, the C2 domain may also be crucial for an efficient targeting of Nedd4 to the plasma membrane. However, it is of interest to observe that in the absence of the C2 domain, Nedd4 can localise, at least in part, to the plasma membrane.

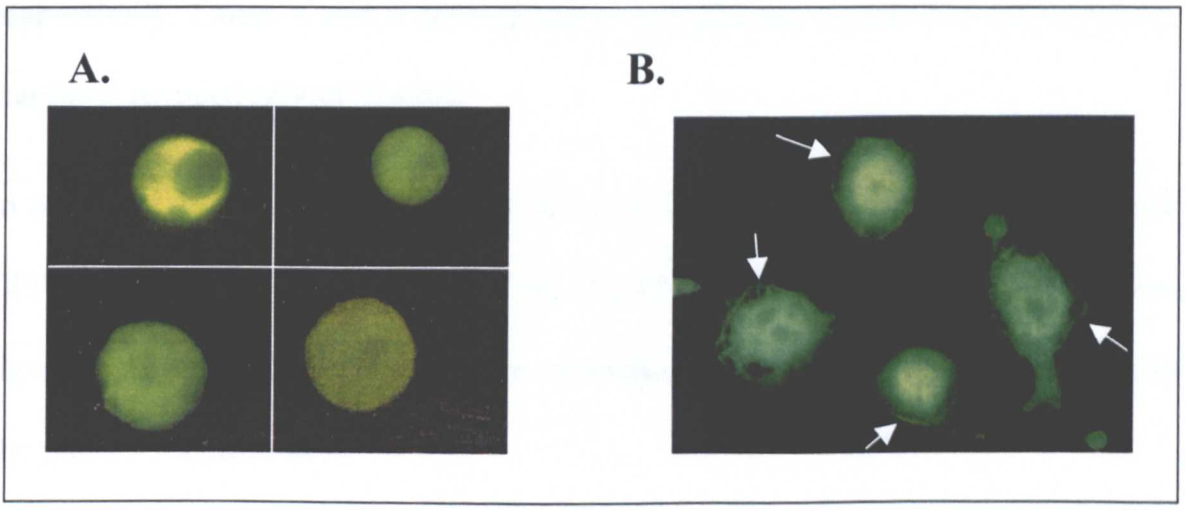

Figure 62: Indirect immunofluorescence analysis of the cellular localisation of hNedd4: (A) WW domains 2, 3 and 4 and (B) WW domains 2, 3 4-HECT.

The cell localisation results were further confirmed by analysing the cytosolic and plasma membrane fractions of 293T cells transfected with hNedd4 WW domains 2,3 and 4 and with hNedd4 WW domains 2, 3 and 4 and the HECT domain (figure 63).

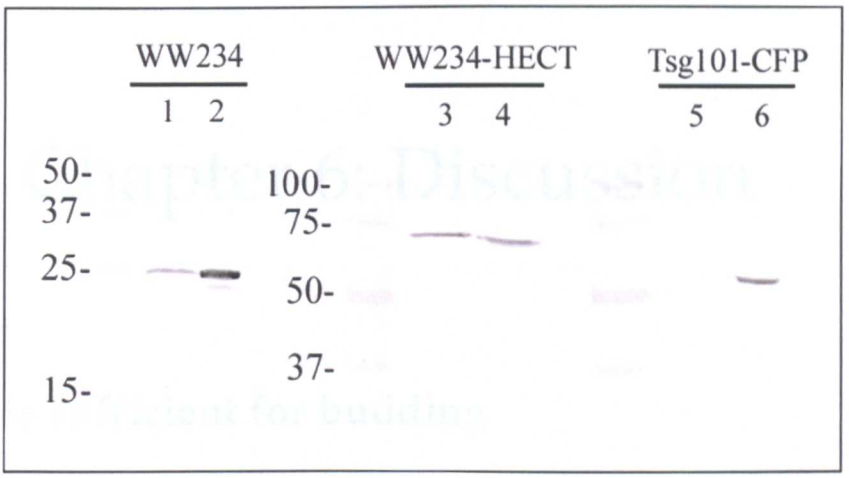

Figure 63: Western blot analysis of the subcellular localisation of human Nedd4 and Tsg101. Lanes 1 and 3 correspond to the plasma membrane fraction of 293T cells transfected with WW234 and WW234-HECT respectively, while lanes 2 and 4 correspond to the cytosolic fractions of WW234 and WW234-HECT respectively. Lanes 5 and 6 correspond to the plasma membrane and cytosolic fractions respectively of Tsg101.

In addition, 293T cells were transfected with Tsg101-CFP and a monoclonal anti-GFP antibody was used to detect the Tsg101 fusion protein (figure 63). From figure 63, it is also clear that Tsg101 is cytosolic, since no protein can be detected at the plasma membrane.

Chapter 6: Discussion

1. VP40 is sufficient for budding

1.1 VP40 localises to the plasma membrane *in vivo*

Viral matrix proteins exert their fundamental role in assembly and budding through their interaction with cellular membranes. The *in vitro* binding properties of three soluble forms of the Ebola virus matrix protein VP40 have previously been characterised which showed that (i) the N-terminal 31 residues were dispensable for membrane association; (ii) the C-terminal residues 213 to 326 were absolutely necessary for membrane targeting; (iii) the removal of the C-terminal domain led to oligomerisation of VP40 (31-212) (Ruigrok et al., 2000b); and (iv) the same ring-like structure is found for full-length VP40 when seven C-terminal residues are deleted, and can be induced by urea treatment and liposome binding *in vitro* (Scianimanico et al., 2000). The structure of VP40 is composed of two domains that are only weakly associated with each other and it has been proposed that a movement of the C-terminal domain induces oligomerisation of VP40 upon membrane association (Dessen et al., 2000b; Scianimanico et al., 2000). Here we show that all three VP40 forms show the same membrane binding properties *in vivo* as observed *in vitro* (Ruigrok et al., 2000b): membrane association absolutely requires the C-terminal domain. Only full-length VP40 and VP40 (31-326) show a clear plasma membrane localisation. Nonetheless, only a relatively small percentage of VP40, VP40 (1-326) as well as VP40 (31-326), is found associated

with the plasma membrane in mammalian cells, using our plasma membrane isolation technique (see Materials and Methods). Interestingly, more recently Kolesnikova *et al.* investigated the cellular localisation of Marburg virus VP40 in Marburg virus-infected cells and in cells expressing recombinant VP40 (Kolesnikova et al., 2002). In both cases, Marburg VP40 alone was found to be associated with membranes of multivesicular bodies and of the late endosomal compartment, suggesting that VP40 can use the intracellular compartment for sorting and accumulation. The significance of this intracellular membrane association in terms of virus assembly is not clear.

1.2 Release of virus-like particles

In addition, our data suggest that Ebola VP40, when associated with the plasma membrane, is released into the cell culture supernatant. Sucrose floatation experiments with the cell culture supernatant imply that VP40 is associated with lipid vesicles. The protection from trypsin digestion further suggests that VP40 is inside the vesicles. Negative staining electron microscopy provides additional evidence that vesicles, which bud off the plasma membrane, resemble virus-like structures (Timmins et al., 2001). However, we only detect virus-like particles at a very low efficiency, which probably underlines an important role of other viral proteins in particle formation (Jin et al., 1997; Sanderson et al., 1994; Schmitt et al., 1999; Schnell et al., 1998). Filovirus show a uniform diameter of approximately 80 nm with some variation in length (Peters, 1996). These measurements are in agreement with the sizes of the observed filamentous particles released from 293T cells expressing full-length VP40 (1-326) (Timmins et al., 2001) or VP40 (31-326) (data not shown). Overall, our experiments show that VP40 is able to assemble

into particles, most likely at the plasma membrane, which can bud off the membrane. The resulting particles are probably internally lined with VP40, in a process that can be achieved in the absence of any other viral protein.

Recently further evidence has been put forward for release of filovirus-like particles by transfected cells expressing either VP40 alone (Jasenosky et al., 2001) or in combination with the glycoprotein GP (Bavari et al., 2002). The efficiency of virus particle release appears to be improved in the presence of the glycoprotein. Moreover, Bavari *et al.* clearly demonstrate that Ebola virus GP specifically localises to detergent-resistant membranes, also known as rafts. Interestingly, we found that the matrix protein, VP40, was not associated with lipid rafts (data not shown), a result which may be due to the fact that VP40 is a peripheral membrane protein and may dissociate during the Triton-X100 extraction. When GP and VP40 were co-expressed, however, both were found to be present in detergent-resistant membranes, suggesting that association of VP40 with GP drives VP40 into the rafts (Bavari et al., 2002). At present, a number of viral structural proteins, and notably transmembrane glycoproteins, have also been reported to associate with lipid rafts. This is the case for influenza virus hemagglutinin and neuraminidase (Scheiffele et al., 1997; Zhang et al., 2000), for a significant proportion of Measles virus (Manie et al., 2000; Vincent et al., 2000) and Sendai virus (Sanderson et al., 1995) structural proteins and for HIV-1 Gag protein (Ono and Freed, 2001). In all these studies, lipid rafts appear to be important for virus assembly and budding, and have been suggested to act as platforms for the stepwise assembly of viruses at the plasma membrane.

1.3 Efficient release requires a full N-terminus of VP40

A similar release of vesicles containing matrix protein has been reported for VSV, which is a member of the rhabdoviruses (order Mononegavirales) (Harty et al., 1999; Justice et al., 1995; Li et al., 1993; Sakaguchi et al., 1999). VSV M also contains a putative WW domain-binding motif (PPxY motif), which has been suggested to bind to cellular factor(s), to enhance the function of M in assembly and particle release (Harty et al., 1999). Ebola virus VP40 has a similar motif at its N-terminus and a recent paper by Harty *et al.* (Harty et al., 2000) shows that the PPxY motif is important for budding and confers VP40's interaction with Rsp5p, a ubiquitin ligase of the Nedd4 family, known to function in endocytosis in yeast and in mammalian cells. Our finding that the deletion of the complete PPxY motif sequence, VP40 (31-326), still supports budding albeit to a five-fold decreased efficiency is in agreement with the mutagenesis data of the PPxY motif which abolished the binding to the ubiquitin ligase and reduced the release of vesicles four-fold (Harty et al., 2000). These results imply that other factors are probably also involved in the budding process. Interestingly, ubiquitin has also been recently implicated to play an important role in the retroviral budding machinery (Patnaik et al., 2000; Strack et al., 2000) indicating a common mechanism for particle release for retroviruses, filoviruses and rhabdoviruses. In addition, we provide evidence that the deletion of the PPxY motif containing sequence does not affect membrane association of VP40 implied by the same percentage of VP40 found to be associated with the plasma membrane *in vivo* for full-length VP40 as well as for VP40 (31-326). These findings are consistent with the data obtained *in vitro* with the liposome binding assays (Ruigrok et al., 2000b).

Although we attempted to determine the oligomeric state of VP40 present in the vesicles released from cells, chemical cross-linking experiments generally resulted in high molecular weight bands whose interpretation was not conclusive (data not shown), and could not be clearly related to the defined oligomeric state observed *in vitro* (Scianimanico et al., 2000). However, the formation of particles driven by the matrix protein indirectly infers the assembly of higher molecular weight structures by VP40, which eventually provide the architecture for particle release. The release of virus-like particles containing VP40 provides evidence that VP40 is able to "self-assemble" on its own into a higher order structure on membranes, a feature which seems to be a prerequisite for membrane-containing particle formation. The efficiency of such a process is dependent on the interaction with cellular protein(s) (Harty et al., 2000; Timmins et al., 2001) and the presence of other viral proteins, such as the glycoprotein (Bavari et al., 2002) will most likely enhance the assembly and budding efficiency. Nevertheless, our results strongly suggest that VP40 contains the minimal information, necessary to induce release of VP40-containing virus-like particles from mammalian cells.

2. Host factors are required for efficient budding

2.1 hNedd4 binds preferentially to oligomeric VP40

Although it is clear that viral matrix proteins play a central role during the assembly and budding of viruses at the cell surface, many questions remain, however, concerning the role of both viral and host proteins in the late stages of the viral life cycle. The binding of VP40 to lipid membranes, both *in vitro* (Ruigrok et al., 2000b) and *in vivo* (Timmins et al., 2001), is dependent on the presence of an

intact C-terminal domain, but does not require a full N-terminus, containing the PPxY and P(T/S)AP motifs. These motifs, however, appear to be required for efficient release of VP40 into the cell culture supernatant of transfected cells (Timmins et al., 2001), suggesting that cellular factors may be involved in protein-protein interactions with VP40 so as to render the assembly and budding processes more efficient. It will now be of interest to see whether the removal of these motifs allows the formation of viruses in a reverse genetics system.

We, therefore, also investigated the possible interactions of VP40 with a member of the growing Nedd4 family of ubiquitin ligases. A recent paper by Harty *et al.* provided the first evidence that the PPxY motif of VP40 is able to mediate interactions with and be ubiquitinated by full-length Rsp5p protein (yeast homologue of mammalian Nedd4), a ubiquitin ligase having multiple WW domains (Harty et al., 2000). However, there was no evidence for any interactions between VP40 and the more relevant mammalian Nedd4 proteins. In the present study we have shown that: (i) VP40 can clearly interact with the WW domains 2, 3 and 4 of human Nedd4, (ii) this interaction requires an intact N-terminus including the PPxY motif, (iii) the oligomeric form of VP40 displays a significantly higher affinity for hNedd4 WW domains, than does monomeric VP40 and finally, (iv) WW domain 3 of hNedd4 is necessary and sufficient for the binding to VP40 *in vitro*. Taken together these results suggest that, although VP40 is sufficient for mediating viral assembly and budding, these events are most certainly facilitated by viral-host interactions.

In this study we were particularly interested in looking at the influence of the oligomeric state of VP40 on its ability to interact with hNedd4. *In vitro* studies have shown that upon membrane binding VP40 undergoes a conformational

change resulting in its oligomerisation (Scianimanico et al., 2000). Electron microscopy analysis of this oligomeric form of VP40 demonstrated that VP40 could form ring-like structures on the membranes, which have been suggested to constitute the building blocks for viral assembly at the host's plasma membrane. Interestingly, we find that the WW domains 2, 3 and 4 of hNedd4 bind to a greater extent to these ring-like structures of VP40, than to the monomeric form of VP40. The biochemical "pull-down" assays clearly demonstrate that VP40 (1-212) can bind to the WW domains of human Nedd4, whereas no detectable interaction can be seen in the case of VP40 (1-326). These results were further confirmed by surface plasmon resonance biosensor studies. This work clearly established that, as in the biochemical studies, VP40 (1-212) could specifically bind to Nedd4 (with and without the HECT domain) and the binding affinity (K_d) was estimated to be close to 1 μM. The dissociation constants for WW domain-PPxY motif complexes lie in the high nanomolar to low millimolar range (Sudol and Hunter, 2000), which is in full agreement with our measurement. In addition, the surface plasmon resonance studies also reveal that VP40 (1-326) can bind to Nedd4, but that this binding is 5-6 fold less efficient than the binding of VP40 (1-212) to Nedd4. This difference may reflect the different behaviour of the monomeric and the oligomeric VP40, as observed in the "pull-down" assays, but may also be due, at least in part, to the differential immobilisation methods. Indeed, VP40 (1-212) was successfully immobilised on an NTA-sensor chip, on which all the molecules are positioned in the same orientation, while VP40 (1-326) was immobilised via its amine groups. In the case of VP40 (1-326) the molecules are, therefore not all positioned in the same orientation, and it cannot be ruled out that the covalent immobilisation may interfere with protein-protein interactions. In addition, VP40

(1-326) is a metastable molecule and it has been observed that a small proportion of the protein can oligomerise over time. This may also account for the weak binding of VP40 (1-326) to Nedd4 in the surface plasmon resonance studies. In both cases, however, the formation and dissociation of the complexes between VP40 and Nedd4 are very rapid, and these complexes are most likely very transient. This may also explain our observation that the complexes between VP40 (1-212) and Nedd4 WW domains are not stable during gel filtration chromatography.

Taken together these results suggest that, *in vivo*, hNedd4 would preferentially interact with the plasma membrane associated VP40, during the late stages of viral infection. This may be explained by a differential accessibility of the N-terminus of VP40 in the monomeric and oligomeric form. The crystal structure of VP40 (monomeric) suggests that the N-terminal part (residues 1-40) may be flexible, since it could be removed by proteolysis with trypsin (Dessen et al., 2000a). However, further proteolysis data also suggests that the intact N-terminus of VP40 can, to a certain extent, protect from further cleavage at lysine 212. This would imply that the N-terminus might take on an extended conformation and extend over to the C-terminal domain, thus protecting lysine 212. The consequence of which would be a reduced accessibility to external factors, such as Nedd4 that bind to the PPxY motif. In contrast, in the oligomeric structure, the N-terminus is easily proteolysed and residues 31-71 are disordered in the structure (unpublished data), strongly suggesting that the N-terminus is more flexible and accessible for binding by Nedd4.

2.2 WW domain 3 of hNedd4 is sufficient for binding to VP40

Nedd4 was originally isolated from a mouse brain library (Kumar et al., 1992). The yeast Rsp5 protein, like its mammalian counterparts from mouse and rat, contains three WW domains, whereas the human Nedd4 homologue contains four WW domains. Based on sequence similarity, the three WW domains of rat and mouse Nedd4 correspond to WW domains 1, 2 and 4 in hNedd4. Interestingly, in this work presented here, we find that VP40 interacts most strongly with WW domain 3 of human Nedd4. WW domain 3, unlike WW domains 2 and 4 is certainly necessary and sufficient for binding to VP40 *in vitro*. Sequence comparison of human Nedd4 WW domains with those of yeast Rsp5 and mouse Nedd4, also reveals that human Nedd4 WW domain 3 is most closely related to the yeast Rsp5 WW domain 2, which can also bind to VP40 (figure 64).

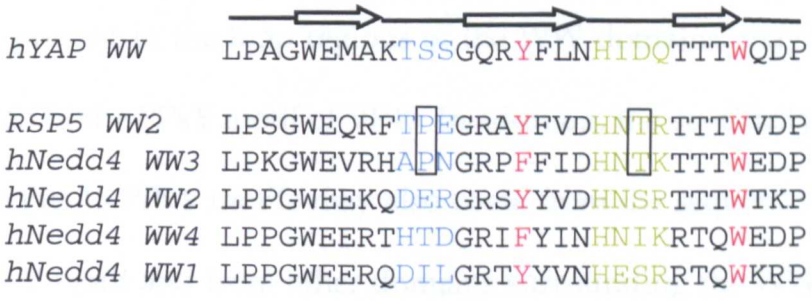

Figure 64: Sequence alignment of the WW domains of human Yes-associated protein (YAP), of Rsp5 (WW domain 2) and of human Nedd4 (WW domains 1, 2, 3 and 4). The secondary structure of human YAP is shown above the sequences, based on the solution structure (Macias et al., 1996). The coloured amino acids are important for the binding to their respective PPxY motifs. The boxed amino acids are conserved in WW domain 2 of Rsp5 and WW domain 3 of hNedd4, but not in the other WW domains of human Nedd4.

A number of amino acid residues are important for recognition of the PPxY motif, as has been determined by the structures of several WW domains in complex with their substrate peptides (Huang et al., 2000; Macias et al., 1996; Macias et al., 2000; Verdecia et al., 2000). These residues are shown in colour in figure 64. The residues in red are highly conserved and are required for the binding of the two proline residues in the PPxY motif. The residues in blue and green, which constitute two loop regions, also appear to be important for the specific recognition of the substrate (Macias et al., 2002). More particularly, the first strand and the loop 1 display the most significant sequence divergences and must, therefore, be responsible to a large extent for the different substrate specificities observed (Macias et al., 2002). Interestingly, Rsp5 WW domain 2 and hNedd4 WW domain 3 share two residues (a proline and a threonine) that are not found in the other three WW domains of human Nedd4, suggesting that these residues, present in the loop regions of the WW domains, may be crucial for the recognition of the PPxY motif of VP40. In addition, the specificity of the binding of WW domains to PPxY motifs may also come from the sequence at the C-terminal of the PPxY motif and from other domains surrounding the WW domain, such as the HECT domain.

These results, however, do not rule out the possibility that WW domains 2 and 4 are involved in the interaction with VP40 *in vivo*. Multiple interactions may be required to improve the overall affinity of VP40 for hNedd4. However, since Ebola virus is highly infectious in mice and rats and that WW domain 3 is absent in rat and mouse Nedd4, this suggests that the pattern of interactions between WW domains and the PPxY motif of VP40 and the affinities of these interactions may not be conserved between species or that other WW domains may also be

involved. Different WW domains from a given Nedd4 protein are thus able to interact with varying affinity with a given substrate, and in addition can each interact individually with different substrates. Moreover, there are now at least fourteen Nedd4-like proteins in humans that probably display different tissue or developmental expression patterns, and this may be an important factor for the regulation of the binding of Nedd4 to substrates such as VP40. Mouse and rat Nedd4 are ubiquitously expressed, but with strongest expression in lung, kidney and brain (Kumar et al., 1997; Staub et al., 1996). Three mRNA species for human Nedd4 were found, which displayed varied expression patterns among normal tissues (Anan et al., 1998).

2.3 Is Nedd4 a central component of viral budding?

Nedd4 proteins have now been implicated in an increasing number of cellular processes (Hicke, 2001), such as internalisation of cell surface receptors (Hicke, 1997; Hicke, 1999; Strous and Govers, 1999), budding of multivesicular bodies into the endosomes (Beck et al., 1999), and finally in virus budding (Patnaik et al., 2000; Strack et al., 2000). Interestingly, the formation of multivesicular bodies in yeast is topologically very similar to virus budding from the plasma membrane; these mechanisms, like other membrane fusion events are energetically unfavourable, and thus require additional factors to improve their efficiency.

There is now growing evidence for the involvement of ubiquitination in viral budding and ubiquitin ligases (E3 enzymes) as host factors that interact with viral proteins to enhance the late assembly and budding processes. A decade ago, avian retroviruses were shown to be able to specifically package free ubiquitin

into their virions at a level 5-fold higher than that found in the cytosol (Putterman et al., 1990). More recently, similar amounts of free ubiquitin were found in HIV-1, SIV and murine leukaemia virus (Ott et al., 1998). Moreover, a number of retrovirus Gag proteins have been shown to be ubiquitinated, a process that requires the presence of a functional late domain, containing a PPxY and/or a P(T/S)AP motif (Patnaik et al., 2000; Strack et al., 2000). In addition, Patnaik *et al* showed that the release of RSV from infected cells is considerably reduced in the presence of proteasome inhibitors, an effect which can be alleviated by overexpressing ubiquitin in *trans* or as a Gag-ubiquitin fusion (Patnaik et al., 2000). These results strongly suggest that ubiquitin is a component of the retrovirus budding machinery. In the case of rhabdo- and filoviruses it is not yet clear whether the matrix proteins become ubiquitinated *in vivo*. Harty *et al.* have shown that Ebola virus VP40 can be ubiquitinated *in vitro* by the yeast Rsp5p protein (Harty et al., 2000). However, so far there is no evidence that the VP40 present in virus-like particles is ubiquitinated. Nevertheless, the finding that human Nedd4 can specifically interact with Ebola VP40, in addition to the evidence for the involvement of other members of the Nedd4 family in the budding mechanisms of rhabdo- and retroviruses, strongly suggests that a common host pathway may be used for efficient release of these viruses from infected cells.

Moreover, Bavari *et al.* have shown that Ebola and Marburg viral proteins are compartmentalised within lipid rafts during viral assembly and budding (Bavari et al., 2002). These lipid rafts appear to support both the entry and exit of filoviruses. Interestingly, Nedd4 has also been found to associate with lipid rafts (Lafont and Simons, 2001; Plant et al., 2000). The presence of the glycoprotein, VP40 and Nedd4 in lipid microdomains may be important for a spatial and

temporal regulation of the budding process. In addition, the presence of these proteins in restricted areas of the plasma membrane may also contribute to an improved assembly and budding.

2.4 VP40 also interacts with human Tsg101

Interestingly, a recent paper from Garrus *et al.* demonstrates that Tsg101, an E2-ubiquitin conjugating enzyme involved in the formation of multivesicular bodies, is required for HIV-1 budding, and suggests that this protein facilitates budding by linking the late domain of HIV-1 Gag protein to the vacuolar protein sorting machinery (Garrus et al., 2001). Tsg101 has an ubiquitin-conjugating-like domain (UEV) that lacks the catalytic cysteine residue (figure 65), and thus is unlikely to catalyse ubiquitination of protein substrates (Babst et al., 2000). Nonetheless, the UEV domain of Tsg101 has been shown to act as a multifunctional domain that can simultaneously bind to both ubiquitin and the P(T/S)AP motif of HIV-1 p6 protein (Garrus et al., 2001).

From the recent three-dimensional structure of the yeast Mms2/UBC13 complex, it appears that UEV domains display an overall similarity to E2-enzyme folds (VanDemark et al., 2001) and the recently solved three-dimensional structure of the UEV domain of Tsg101 also reveals an E2-like fold (Garrus et al., 2001). This structural data supports the biochemical data, suggesting that UEV domains of both Mms2 and Tsg101 can bind simultaneously to ubiquitin and to specific protein-protein recognition motifs, such as the P(T/S)AP motif, resulting in the formation of large multiprotein complexes.

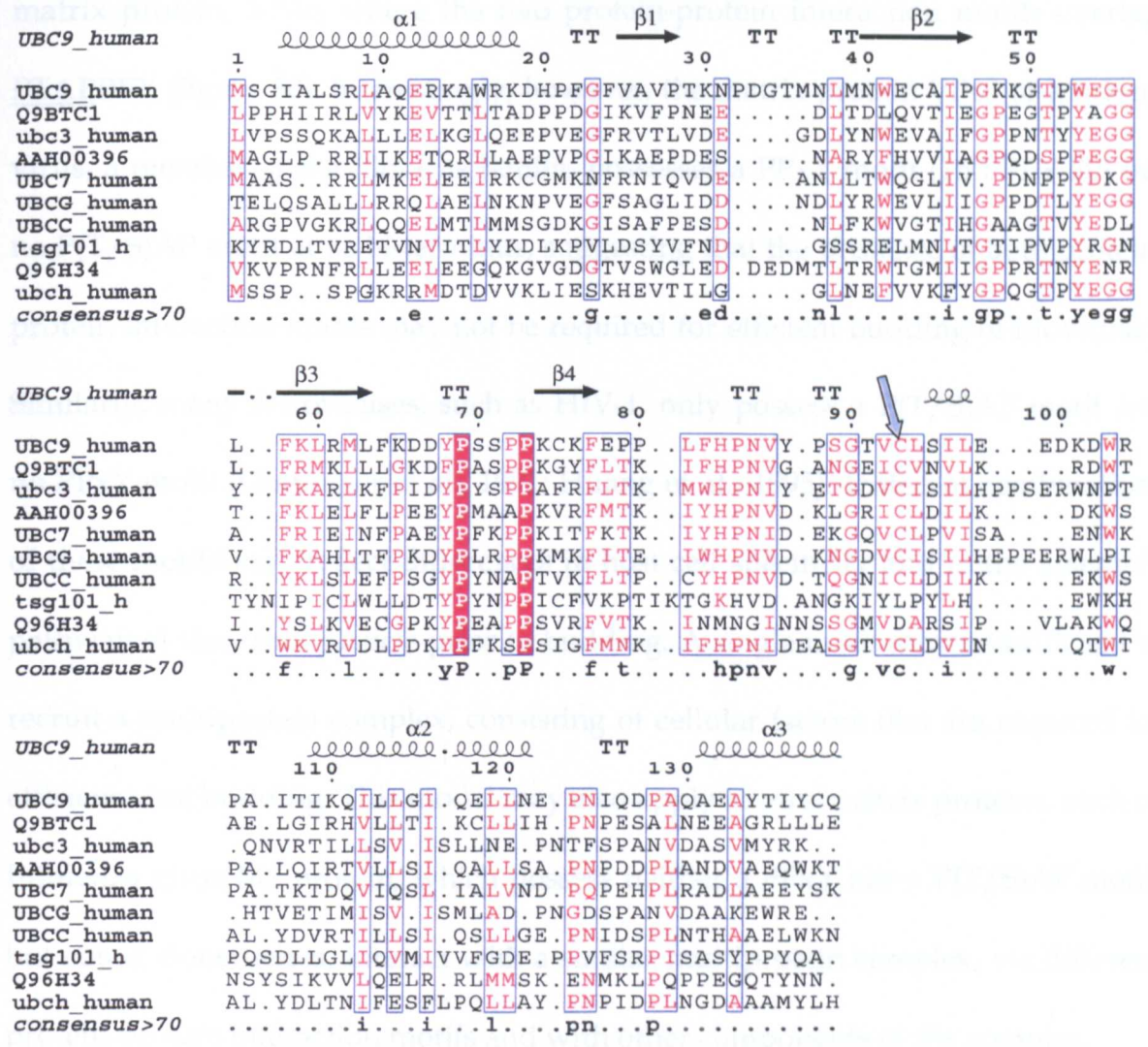

Figure 65: Amino acid sequence alignment of human ubiquitin-conjugating enzymes (E2), showing the conserved residues and the secondary structure of human UBC9, for which the three-dimensional structure is known (PDB ID: 1A3S). Alpha helices are numbered as α1, α2 and α3, beta strands are designated as β1, β2, β3 and β4, and beta turns are labelled as TT. A blue arrow shows the active site cysteine, which is missing in Tsg101.

Tsg101 and Nedd4 may be playing complementary roles in viral budding, since binding sites for both Tsg101 (PT/SAP) and Nedd4 (PPxY) are found in a number of viral proteins (Strack et al., 2000), but also in cellular proteins such as connexins 43 and 45 (Berthoud et al., 2000). This is the case notably for Ebola virus

matrix protein, VP40, where the two protein-protein interaction motifs overlap: PTAPPEY (figure 30). Interestingly, however, the matrix protein VP40 of Marburg virus, a member of the filovirus family, possesses a PPxY motif (15-PPPY-20) but no P(T/S)AP motif at its N-terminus, suggesting that the presence of both protein-protein interaction motifs may not be required for efficient budding of filoviruses. Similarly, many retroviruses, such as HIV-1, only possess a P(T/S)AP motif and no PPxY motif (Gottlinger et al., 1991; Huang et al., 1995). Different combinations of these motifs may reflect differences in host cell specificity and in the assembly pathway of the virus particle prior to budding. A single motif may be sufficient to recruit a multiprotein complex, consisting of cellular factors that are required for efficient viral budding. This model may also apply to viral matrix proteins, such as Influenza virus M1 protein, which possess neither a PPxY nor a P(T/S)AP motif, but could, nonetheless, interact with a similar multiprotein complex, via different protein-protein interaction motifs and with other components of the complex.

In this report, we clearly demonstrate that the first 30 residues of VP40, containing these two motifs, are required not only for binding to hNedd4 WW domains 2, 3 and 4, but also to the N-terminal domain (UEV domain) of human Tsg101, *in vitro*. These results were obtained using a biochemical "pull-down" assay and surface plasmon resonance studies. The latter also showed that the complex formation between Tsg101 and VP40 was rapid (as in the case of Nedd4), while the dissociation appeared to be slower, suggesting that the complex may be more stable. Unfortunately, no reliable binding constants could be obtained. In addition, the biochemical data revealed that Tsg101, unlike Nedd4, displayed no preference for binding to oligomeric VP40 versus monomeric VP40. These results suggest that Tsg101 and Nedd4 may be binding to VP40 at distinct stages of viral

assembly and budding. Martin-Serrano *et al.* have recently shown that Ebola VP40 recruits Tsg101 to the plasma membrane, where viral assembly then occurs (Martin-Serrano et al., 2001). Together with our results, this suggests that cytosolic Tsg101 binds to VP40 in the cytoplasm (monomeric VP40) and then targets VP40 to the site of particle formation, where VP40 may interact with both the lipid membrane, leading to oligomerisation and with further cellular factors, such as Nedd4. Interestingly, we find that Nedd4 is able to localise to the plasma membrane in the absence of its C2 domain. The fact that in VP40 the two binding motifs overlap, strongly suggest that the binding to hNedd4 and hTsg101 *in vivo* must be sequential.

More recently, Kolesnikova *et al.* (Kolesnikova et al., 2002) investigated the cellular localisation of Marburg virus VP40. In this study, VP40 was found to be, in part, associated with multivesicular bodies (MVBs) in Marburg virus-infected cells and in HeLa cells expressing VP40 alone. These results constitute further evidence for the possible involvement of the vacuolar protein sorting machinery in viral assembly and budding. Since Marburg VP40 does not possess a P(T/S)AP motif for binding to Tsg101, the targeting of VP40 to MVBs must be dependent on an additional host factor. Thus, a search for new cellular factors involved in virus budding may reveal new hints, at a molecular level, as to how viruses use cellular pathways to complete their life cycle.

In addition, the binding of either hNedd4 or hTsg101 to VP40 resulted in the stabilisation of the N-terminus of VP40, which became less sensitive to proteolytic degradation during the initial purification steps and when stored at 4°C for several days. Preliminary crystallisation trials of VP40 (1-212) in complex with either WW domain 3 of hNedd4 or the N-terminal domain of Tsg101 are

currently being carried out. This structural information will improve our understanding of the protein-protein contacts involved in the specific recognition of the different oligomeric forms of VP40 by hNedd4 and hTsg101. This structural approach, together with further studies on the role of hNedd4 and hTsg101 in filovirus assembly and budding *in vivo,* will most certainly provide us with a clearer picture of the host-pathogen interactions involved in the late stages of the virus life cycle. So far, the finding that both Nedd4-like proteins and Tsg101, a member of the vacuolar protein sorting machinery are involved in the assembly and budding of retroviruses (HIV-1) and filoviruses (Ebola), strongly suggests that enveloped viruses may use a common cellular pathway for efficient budding of viral particles.

References:

Anan, T., Nagata, Y., Koga, H., Honda, Y., Yabuki, N., Miyamoto, C., Kuwano, A., Matsuda, I., Endo, F., Saya, H., and Nakao, M. (1998). Human ubiquitin-protein ligase Nedd4: expression, subcellular localization and selective interaction with ubiquitin-conjugating enzymes. Genes Cells *3*, 751-763.

Babst, M., Odorizzi, G., Estepa, E. J., and Emr, S. D. (2000). Mammalian tumor susceptibility gene 101 (TSG101) and the yeast homologue, Vps23p, both function in late endosomal trafficking. Traffic *1*, 248-258.

Baize, S., Leroy, E. M., Georges-Courbot, M. C., Capron, M., Lansoud-Soukate, J., Debre, P., Fisher-Hoch, S. P., McCormick, J. B., and Georges, A. J. (1999). Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. Nat Med *5*, 423-426.

Baize, S., Leroy, E. M., Mavoungou, E., and Fisher-Hoch, S. P. (2000). Apoptosis in fatal Ebola infection. Does the virus toll the bell for immune system? Apoptosis *5*, 5-7.

Barklis, E., McDermott, J., Wilkens, S., Fuller, S., and Thompson, D. (1998). Organization of HIV-1 capsid proteins on a lipid monolayer. J Biol Chem *273*, 7177-7180.

Baskerville, A., Fisher-Hoch, S. P., Neild, G. H., and Dowsett, A. B. (1985). Ultrastructural pathology of experimental Ebola haemorrhagic fever virus infection. J Pathol *147*, 199-209.

Bavari, S., Bosio, C. M., Wiegand, E., Ruthel, G., Will, A. B., Geisbert, T. W., Hevey, M., Schmaljohn, C., Schmaljohn, A., and Aman, M. J. (2002). Lipid raft microdomains: a gateway for compartmentalized trafficking of ebola and marburg viruses. J Exp Med *195*, 593-602.

Beck, T., Schmidt, A., and Hall, M. N. (1999). Starvation induces vacuolar targeting and degradation of the tryptophan permease in yeast. J Cell Biol *146*, 1227-1238.

Becker, S., Spiess, M., and Klenk, H. D. (1995). The asialoglycoprotein receptor is a potential liver-specific receptor for Marburg virus. J Gen Virol *76*, 393-399.

Berger, E. A., Murphy, P. M., and Farber, J. M. (1999). Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. Annu Rev Immunol *17*, 657-700.

Berthoud, V. M., Tadros, P. N., and Beyer, E. C. (2000). Connexin and gap junction degradation. Methods *20*, 180-187.

Bork, P., and Sudol, M. (1994). The WW domain: a signalling site in dystrophin? Trends Biochem Sci *19*, 531-533.

Buechi, M., and Bachi, T. (1982). Microscopy of internal structures of Sendai virus associated with the cytoplasmic surface of host membranes. Virology *120*, 349-359.

Chan, S. Y., Speck, R. F., Ma, M. C., and Goldsmith, M. A. (2000). Distinct mechanisms of entry by envelope glycoproteins of Marburg and Ebola (Zaire) viruses. J Virol *74*, 4933-4937.

Chan, S. Y., Empig, C. J., Welte, F. J., Speck, R. F., Schmaljohn, A., Kreisberg, J. F., and Goldsmith, M. A. (2001). Folate receptor-alpha is a cofactor for cellular entry by Marburg and Ebola viruses. Cell *106*, 117-126.

Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., and Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. Science *243*, 1576-1583.

Chen, H. I., and Sudol, M. (1995). The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. Proc Natl Acad Sci U S A *92*, 7819-7823.

Chong, L. D., and Rose, J. K. (1993). Membrane association of functional vesicular stomatitis virus matrix protein in vivo. J Virol *67*, 407-414.

Christensen, A. M., Massiah, M. A., Turner, B. G., Sundquist, W. I., and Summers, M. F. (1996). Three-dimensional structure of the HTLV-II matrix protein and comparative analysis of matrix proteins from the different classes of pathogenic human retroviruses. J Mol Biol *264*, 1117-1131.

Conte, M. R., Klikova, M., Hunter, E., Ruml, T., and Matthews, S. (1997). The three-dimensional solution structure of the matrix protein from the type D retrovirus, the Mason-Pfizer monkey virus, and implications for the morphology of retroviral assembly. Embo J *16*, 5819-5826.

Conte, M. R., and Matthews, S. (1998). Retroviral matrix proteins: a structural perspective. Virology *246*, 191-198.

Coronel, E. C., Murti, K. G., Takimoto, T., and Portner, A. (1999). Human parainfluenza virus type 1 matrix and nucleoprotein genes transiently expressed in mammalian cells induce the release of virus- like particles containing nucleocapsid-like structures. J Virol *73*, 7035-7038.

Dessen, A., Forest, E., Volchkov, V., Dolnik, O., Klenk, H. D., and Weissenhorn, W. (2000a). Crystallization and preliminary X-ray analysis of the matrix protein from Ebola virus. Acta Crystallogr D Biol Crystallogr *56*, 758-760.

Dessen, A., Volchkov, V., Dolnik, O., Klenk, H. D., and Weissenhorn, W. (2000b). Crystal structure of the matrix protein VP40 from Ebola virus. Embo J *19*, 4228-4236.

Dowell, S. F., Mukunu, R., Ksiazek, T. G., Khan, A. S., Rollin, P. E., and Peters, C. J. (1999). Transmission of Ebola hemorrhagic fever: a study of risk factors in family members, Kikwit, Democratic Republic of the Congo, 1995. Commission de Lutte contre les Epidemies a Kikwit. J Infect Dis *179 Suppl 1*, S87-91.

Dunn, R., and Hicke, L. (2001). Domains of the Rsp5 ubiquitin-protein ligase required for receptor-mediated and fluid-phase endocytosis. Mol Biol Cell *12*, 421-435.

Dupre, S., Volland, C., and Haguenauer-Tsapis, R. (2001). Membrane transport: ubiquitylation in endosomal sorting. Curr Biol *11*, R932-934.

Elliott, L. H., Kiley, M. P., and McCormick, J. B. (1985). Descriptive analysis of Ebola virus proteins. Virology *147*, 169-176.

Enami, M., and Enami, K. (1996). Influenza virus hemagglutinin and neuraminidase glycoproteins stimulate the membrane association of the matrix protein. J Virol *70*, 6653-6657.

Farr, T. J., Coddington-Lawson, S. J., Snyder, P. M., and McDonald, F. J. (2000). Human Nedd4 interacts with the human epithelial Na+ channel: WW3 but not WW1 binds to Na+-channel subunits. Biochem J *345 Pt 3*, 503-509.

Feldmann, H., Volchkov, V. E., Volchkova, V. A., and Klenk, H. D. (1999). The glycoproteins of Marburg and Ebola virus and their potential roles in pathogenesis. Arch Virol Suppl *15*, 159-169.

Fisher-Hoch, S. P., Platt, G. S., Neild, G. H., Southee, T., Baskerville, A., Raymond, R. T., Lloyd, G., and Simpson, D. I. (1985). Pathophysiology of shock and hemorrhage in a fulminating viral infection (Ebola). J Infect Dis *152*, 887-894.

Fisher-Hoch, S. P., Brammer, T. L., Trappier, S. G., Hutwagner, L. C., Farrar, B. B., Ruo, S. L., Brown, B. G., Hermann, L. M., Perez-Oronoz, G. I., Goldsmith, C. S., and et al. (1992). Pathogenic potential of filoviruses: role of geographic origin of primate host and virus strain. J Infect Dis *166*, 753-763.

Gajewska, B., Kaminska, J., Jesionowska, A., Martin, N. C., Hopper, A. K., and Zoladek, T. (2001). WW domains of Rsp5p define different functions: determination of roles in fluid phase and uracil permease endocytosis in Saccharomyces cerevisiae. Genetics *157*, 91-101.

Galan, J. M., Moreau, V., Andre, B., Volland, C., and Haguenauer-Tsapis, R. (1996). Ubiquitination mediated by the Npi1p/Rsp5p ubiquitin-protein ligase is required for endocytosis of the yeast uracil permease. J Biol Chem *271*, 10946-10952.

Garnier, L., Wills, J. W., Verderame, M. F., and Sudol, M. (1996). WW domains and retrovirus budding. Nature *381*, 744-745.

Garoff, H., Hewson, R., and Opstelten, D. J. (1998). Virus maturation by budding. Microbiol Mol Biol Rev *62*, 1171-1190.

Garrus, J. E., von Schwedler, U. K., Pornillos, O. W., Morham, S. G., Zavitz, K. H., Wang, H. E., Wettstein, D. A., Stray, K. M., Cote, M., Rich, R. L., Myszka, D. G., and Sundquist, W. I. (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell *107*, 55-65.

Gaudin, Y., Sturgis, J., Doumith, M., Barge, A., Robert, B., and Ruigrok, R. W. (1997). Conformational flexibility and polymerization of vesicular stomatitis virus matrix protein. J Mol Biol *274*, 816-825.

Geisbert, T. W., and Jahrling, P. B. (1995). Differentiation of filoviruses by electron microscopy. Virus Res *39*, 129-150.

Gilligan, K. J., Geisbert, J.B., Jahrling, P.B. and Anderson, K. (1997). Assessment of protective immunity conferred by recombinant vaccinia viruses to guinea pigs challenged with Ebola virus. In Vaccines 97, F. Brown, Burton, D., Doherty, J., Mekalanos, J. and Norrby, E., ed. (Col Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press), pp. 87-92.

Gilpin, B. J., Loechel, F., Mattei, M. G., Engvall, E., Albrechtsen, R., and Wewer, U. M. (1998). A novel, secreted form of human ADAM 12 (meltrin alpha) provokes myogenesis in vivo. J Biol Chem *273*, 157-166.

Gonzalez, S. A., Affranchino, J. L., Gelderblom, H. R., and Burny, A. (1993). Assembly of the matrix protein of simian immunodeficiency virus into virus-like particles. Virology *194*, 548-556.

Gottlinger, H. G., Sodroski, J. G., and Haseltine, W. A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. Proc Natl Acad Sci U S A *86*, 5781-5785.

Gottlinger, H. G., Dorfman, T., Sodroski, J. G., and Haseltine, W. A. (1991). Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. Proc Natl Acad Sci U S A *88*, 3195-3199.

Goulet, C. C., Volk, K. A., Adams, C. M., Prince, L. S., Stokes, J. B., and Snyder, P. M. (1998). Inhibition of the epithelial Na+ channel by interaction of Nedd4 with a PY motif deleted in Liddle's syndrome. J Biol Chem *273*, 30012-30017.

Harty, R. N., Paragas, J., Sudol, M., and Palese, P. (1999). A proline-rich motif within the matrix protein of vesicular stomatitis virus and rabies virus interacts with WW domains of cellular proteins: implications for viral budding. J Virol *73*, 2921-2929.

Harty, R. N., Brown, M. E., Wang, G., Huibregtse, J., and Hayes, F. P. (2000). A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding. Proc Natl Acad Sci U S A *97*, 13871-13876.

Harty, R. N., Brown, M. E., McGettigan, J. P., Wang, G., Jayakar, H. R., Huibregtse, J. M., Whitt, M. A., and Schnell, M. J. (2001). Rhabdoviruses and the cellular ubiquitin-proteasome system: a budding interaction. J Virol *75*, 10623-10629.

Harvey, K. F., Dinudom, A., Komwatana, P., Jolliffe, C. N., Day, M. L., Parasivam, G., Cook, D. I., and Kumar, S. (1999). All three WW domains of murine Nedd4 are involved in the regulation of epithelial sodium channels by intracellular Na+. J Biol Chem *274*, 12525-12530.

Harvey, K. F., and Kumar, S. (1999). Nedd4-like proteins: an emerging family of ubiquitin-protein ligases implicated in diverse cellular functions. Trends Cell Biol *9*, 166-169.

Heggeness, M. H., Smith, P. R., and Choppin, P. W. (1982). In vitro assembly of the nonglycosylated membrane protein (M) of Sendai virus. Proc Natl Acad Sci U S A *79*, 6232-6236.

Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. Annu Rev Biochem *67*, 425-479.

Hicke, L. (1997). Ubiquitin-dependent internalization and down-regulation of plasma membrane proteins. Faseb J *11*, 1215-1226.

Hicke, L. (1999). Gettin' down with ubiquitin: turning off cell-surface receptors, transporters and channels. Trends Cell Biol *9*, 107-112.

Hicke, L. (2001). Protein regulation by monoubiquitin. Nat Rev Mol Cell Biol *2*, 195-201.

Hill, C. P., Worthylake, D., Bancroft, D. P., Christensen, A. M., and Sundquist, W. I. (1996). Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. Proc Natl Acad Sci U S A *93*, 3099-3104.

Hochstrasser, M. (1996). Ubiquitin-dependent protein degradation. Annu Rev Genet *30*, 405-439.

Hofmann, K., and Falquet, L. (2001). A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems. Trends Biochem Sci *26*, 347-350.

Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P. M., Huibregtse, J. M., and Pavletich, N. P. (1999). Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. Science *286*, 1321-1326.

Huang, M., Orenstein, J. M., Martin, M. A., and Freed, E. O. (1995). p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. J Virol *69*, 6810-6818.

Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M. J. (2000). Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. Nat Struct Biol *7*, 634-638.

Huibregtse, J. M., Scheffner, M., Beaudenon, S., and Howley, P. M. (1995). A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. Proc Natl Acad Sci U S A *92*, 2563-2567.

Ikeda, M., Ikeda, A., Longan, L. C., and Longnecker, R. (2000). The Epstein-Barr virus latent membrane protein 2A PY motif recruits WW domain-containing ubiquitin-protein ligases. Virology *268*, 178-191.

Imhof, M. O., and McDonnell, D. P. (1996). Yeast RSP5 and its human homolog hRPF1 potentiate hormone-dependent activation of transcription by human progesterone and glucocorticoid receptors. Mol Cell Biol *16*, 2594-2605.

Jahrling, P. B., Geisbert, T. W., Dalgard, D. W., Johnson, E. D., Ksiazek, T. G., Hall, W. C., and Peters, C. J. (1990). Preliminary report: isolation of Ebola virus from monkeys imported to USA. Lancet *335*, 502-505.

Jasenosky, L. D., Neumann, G., Lukashevich, I., and Kawaoka, Y. (2001). Ebola virus VP40-induced particle formation and association with the lipid bilayer. J Virol *75*, 5205-5214.

Jin, H., Leser, G. P., Zhang, J., and Lamb, R. A. (1997). Influenza virus hemagglutinin and neuraminidase cytoplasmic tails control particle shape. Embo J *16*, 1236-1247.

Johnson, K. M., Lange, J. V., Webb, P. A., and Murphy, F. A. (1977). Isolation and partial characterisation of a new virus causing acute haemorrhagic fever in Zaire. Lancet *1*, 569-571.

Justice, P. A., Sun, W., Li, Y., Ye, Z., Grigera, P. R., and Wagner, R. R. (1995). Membrane vesiculation function and exocytosis of wild-type and mutant matrix proteins of vesicular stomatitis virus. J Virol *69*, 3156-3160.

Kanda, T. (1996). A ubiquitin-protein ligase (E3) mutation of Saccharomyces cerevisiae suppressed by co-overexpression of two ubiquitin-specific processing proteases. Genes Genet Syst *71*, 75-83.

Kanelis, V., Rotin, D., and Forman-Kay, J. D. (2001). Solution structure of a Nedd4 WW domain-ENaC peptide complex. Nat Struct Biol *8*, 407-412.

Katzmann, D. J., Babst, M., and Emr, S. D. (2001). Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. Cell *106*, 145-155.

Kikonyogo, A., Bouamr, F., Vana, M. L., Xiang, Y., Aiyar, A., Carter, C., and Leis, J. (2001). Proteins related to the Nedd4 family of ubiquitin protein ligases interact with the L domain of Rous sarcoma virus and are required for gag budding from cells. Proc Natl Acad Sci U S A *98*, 11199-11204.

Kiley, M. P., Bowen, E. T., Eddy, G. A., Isaacson, M., Johnson, K. M., McCormick, J. B., Murphy, F. A., Pattyn, S. R., Peters, D., Prozesky, O. W., Regnery, R. L., Simpson, D. I., Slenczka, W.,

Sureau, P., van der Groen, G., Webb, P. A., and Wulff, H. (1982). Filoviridae: a taxonomic home for Marburg and Ebola viruses? Intervirology *18*, 24-32.

Kiley, M. P., Cox, N. J., Elliott, L. H., Sanchez, A., DeFries, R., Buchmeier, M. J., Richman, D. D., and McCormick, J. B. (1988). Physicochemical properties of Marburg virus: evidence for three distinct virus strains and their relationship to Ebola virus. J Gen Virol *69*, 1957-1967.

Klenk, H. D., Volchkov, V. E., and Feldmann, H. (1998). Two strings to the bow of Ebola virus. Nat Med *4*, 388-389.

Kolesnikova, L., Bugany, H., Klenk, H. D., and Becker, S. (2002). VP40, the matrix protein of Marburg virus, is associated with membranes of the late endosomal compartment. J Virol *76*, 1825-1838.

Kretzschmar, E., Bui, M., and Rose, J. K. (1996). Membrane association of influenza virus matrix protein does not require specific hydrophobic domains or the viral glycoproteins. Virology *220*, 37-45.

Kumar, S., Tomooka, Y., and Noda, M. (1992). Identification of a set of genes with developmentally down-regulated expression in the mouse brain. Biochem Biophys Res Commun *185*, 1155-1161.

Kumar, S., Harvey, K. F., Kinoshita, M., Copeland, N. G., Noda, M., and Jenkins, N. A. (1997). cDNA cloning, expression analysis, and mapping of the mouse Nedd4 gene. Genomics *40*, 435-443.

Lafont, F., and Simons, K. (2001). Raft-partitioning of the ubiquitin ligases Cbl and Nedd4 upon IgE- triggered cell signaling. Proc Natl Acad Sci U S A *98*, 3180-3184.

Lemmon, S. K., and Traub, L. M. (2000). Sorting in the endosomal system in yeast and animal cells. Curr Opin Cell Biol *12*, 457-466.

Li, L., Liao, J., Ruland, J., Mak, T. W., and Cohen, S. N. (2001). A TSG101/MDM2 regulatory loop modulates MDM2 degradation and MDM2/p53 feedback control. Proc Natl Acad Sci U S A *98*, 1619-1624.

Li, L. and Cohen, S.N. (1996). Tsg101: a novel tumor susceptibility gene isolated by controlled homozygous functional knockout of allelic loci in mammalian cells. Cell *85*, 319-329.

Li, Y., Luo, L., Schubert, M., Wagner, R. R., and Kang, C. Y. (1993). Viral liposomes released from insect cells infected with recombinant baculovirus expressing the matrix protein of vesicular stomatitis virus. J Virol *67*, 4415-4420.

Luban, J. (2001). HIV-1 and Ebola virus: the getaway driver nabbed. Nat Med *7*, 1278-1280.

Macias, M. J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996). Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. Nature *382*, 646-649.

Macias, M. J., Gervais, V., Civera, C., and Oschkinat, H. (2000). Structural analysis of WW domains and design of a WW prototype. Nat Struct Biol *7*, 375-379.

Macias, M. J., Wiesner, S., and Sudol, M. (2002). WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. FEBS Lett *513*, 30-37.

Manie, S. N., Debreyne, S., Vincent, S., and Gerlier, D. (2000). Measles virus structural components are enriched into lipid raft microdomains: a potential cellular location for virus assembly. J Virol *74*, 305-311.

Martini, G. A. and Siegert, R. (1971). Marburg virus disease, 1 edn (Berlin, Heidelberg, New York, Springer).

Martin-Serrano, J., Zang, T., and Bieniasz, P. D. (2001). HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. Nat Med *7*, 1313-1319.

Matthews, S., Mikhailov, M., Burny, A., and Roy, P. (1996). The solution structure of the bovine leukaemia virus matrix protein and similarity with lentiviral matrix proteins. Embo J *15*, 3267-3274.

McCreedy, B. J., Jr., McKinnon, K. P., and Lyles, D. S. (1990). Solubility of vesicular stomatitis virus M protein in the cytosol of infected cells or isolated from virions. J Virol *64*, 902-906.

Mebatsion, T., Weiland, F., and Conzelmann, K. K. (1999). Matrix protein of rabies virus is responsible for the assembly and budding of bullet-shaped particles and interacts with the transmembrane spike glycoprotein G. J Virol *73*, 242-250.

Monath, T. P. (1999). Ecology of Marburg and Ebola viruses: speculations and directions for future research. J Infect Dis *179 Suppl 1*, S127-138.

Ono, A., and Freed, E. O. (2001). Plasma membrane rafts play a critical role in HIV-1 assembly and release. Proc Natl Acad Sci U S A *98*, 13925-13930.

Ott, D. E., Coren, L. V., Copeland, T. D., Kane, B. P., Johnson, D. G., Sowder, R. C., 2nd, Yoshinaka, Y., Oroszlan, S., Arthur, L. O., and Henderson, L. E. (1998). Ubiquitin is covalently attached to the p6Gag proteins of human immunodeficiency virus type 1 and simian immunodeficiency virus and to the p12Gag protein of Moloney murine leukemia virus. J Virol *72*, 2962-2968.

Patnaik, A., Chau, V., and Wills, J. W. (2000). Ubiquitin is part of the retrovirus budding machinery. Proc Natl Acad Sci U S A *97*, 13069-13074.

Perez, O. D., and Nolan, G. P. (2001). Resistance is futile: assimilation of cellular machinery by HIV-1. Immunity *15*, 687-690.

Peters, C. J., Sanchez, A., Rollin, P.E., Ksiazek, T.G. and Murphy, F.A. (1996). Filoviridae: Marburg and Ebola Viruses. In Fields Virology, B. N. Fields, ed. (Philadelphia, Lippincott-Raven), pp. 1161-1176.

Pickart, C. M. (2001). Ubiquitin enters the new millennium. Mol Cell *8*, 499-504.

Plant, P. J., Lafont, F., Lecat, S., Verkade, P., Simons, K., and Rotin, D. (2000). Apical membrane targeting of Nedd4 is mediated by an association of its C2 domain with annexin XIIIb. J Cell Biol *149*, 1473-1484.

Pushko, P., Parker, M., Ludwig, G. V., Davis, N. L., Johnston, R. E., and Smith, J. F. (1997). Replicon-helper systems from attenuated Venezuelan equine encephalitis virus: expression of heterologous genes in vitro and immunization against heterologous pathogens in vivo. Virology *239*, 389-401.

Pushko, P., Bray, M., Ludwig, G. V., Parker, M., Schmaljohn, A., Sanchez, A., Jahrling, P. B., and Smith, J. F. (2000). Recombinant RNA replicons derived from attenuated Venezuelan equine

encephalitis virus protect guinea pigs and mice from Ebola hemorrhagic fever virus. Vaccine *19*, 142-153.

Putterman, D., Pepinsky, R. B., and Vogt, V. M. (1990). Ubiquitin in avian leukosis virus particles. Virology *176*, 633-637.

Rao, Z., Belyaev, A. S., Fry, E., Roy, P., Jones, I. M., and Stuart, D. I. (1995). Crystal structure of SIV matrix antigen and implications for virus assembly. Nature *378*, 743-747.

Rizo, J., and Sudhof, T. C. (1998). C2-domains, structure and function of a universal Ca2+-binding domain. J Biol Chem *273*, 15879-15882.

Rotin, D., Staub, O., and Haguenauer-Tsapis, R. (2000). Ubiquitination and endocytosis of plasma membrane proteins: role of Nedd4/Rsp5p family of ubiquitin-protein ligases. J Membr Biol *176*, 1-17.

Ruigrok, R. W., Barge, A., Durrer, P., Brunner, J., Ma, K., and Whittaker, G. R. (2000a). Membrane interaction of influenza virus M1 protein. Virology *267*, 289-298.

Ruigrok, R. W., Schoehn, G., Dessen, A., Forest, E., Volchkov, V., Dolnik, O., Klenk, H. D., and Weissenhorn, W. (2000b). Structural characterization and membrane binding properties of the matrix protein VP40 of Ebola virus. J Mol Biol *300*, 103-112.

Sakaguchi, T., Uchiyama, T., Fujii, Y., Kiyotani, K., Kato, A., Nagai, Y., Kawai, A., and Yoshida, T. (1999). Double-layered membrane vesicles released from mammalian cells infected with Sendai virus expressing the matrix protein of vesicular stomatitis virus. Virology *263*, 230-243.

Samaranayake, L. P., Peiris, J. S., and Scully, C. (1996). Ebola virus infection: an overview. Br Dent J *180*, 264-266.

Sanchez, A., Trappier, S. G., Mahy, B. W., Peters, C. J., and Nichol, S. T. (1996). The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. Proc Natl Acad Sci U S A *93*, 3602-3607.

Sanchez, A., Yang, Z. Y., Xu, L., Nabel, G. J., Crews, T., and Peters, C. J. (1998). Biochemical analysis of the secreted and virion glycoproteins of Ebola virus. J Virol *72*, 6442-6447.

Sanderson, C. M., McQueen, N. L., and Nayak, D. P. (1993). Sendai virus assembly: M protein binds to viral glycoproteins in transit through the secretory pathway. J Virol *67*, 651-663.

Sanderson, C. M., Wu, H. H., and Nayak, D. P. (1994). Sendai virus M protein binds independently to either the F or the HN glycoprotein in vivo. J Virol *68*, 69-76.

Sanderson, C. M., Avalos, R., Kundu, A., and Nayak, D. P. (1995). Interaction of Sendai viral F, HN, and M proteins with host cytoskeletal and lipid components in Sendai virus-infected BHK cells. Virology *209*, 701-707.

Scheiffele, P., Roth, M. G., and Simons, K. (1997). Interaction of influenza virus haemagglutinin with sphingolipid- cholesterol membrane domains via its transmembrane domain. Embo J *16*, 5501-5508.

Schmitt, A. P., He, B., and Lamb, R. A. (1999). Involvement of the cytoplasmic domain of the hemagglutinin- neuraminidase protein in assembly of the paramyxovirus simian virus 5. J Virol *73*, 8703-8712.

Schnell, M. J., Buonocore, L., Boritz, E., Ghosh, H. P., Chernish, R., and Rose, J. K. (1998). Requirement for a non-specific glycoprotein cytoplasmic domain sequence to drive efficient budding of vesicular stomatitis virus. Embo J *17*, 1289-1296.

Schubert, U., Ott, D. E., Chertova, E. N., Welker, R., Tessmer, U., Princiotta, M. F., Bennink, J. R., Krausslich, H. G., and Yewdell, J. W. (2000). Proteasome inhibition interferes with gag polyprotein processing, release, and maturation of HIV-1 and HIV-2. Proc Natl Acad Sci U S A *97*, 13057-13062.

Scianimanico, S., Schoehn, G., Timmins, J., Ruigrok, R. H., Klenk, H. D., and Weissenhorn, W. (2000). Membrane association induces a conformational change in the Ebola virus matrix protein. Embo J *19*, 6732-6741.

Sha, B., and Luo, M. (1997). Structure of a bifunctional membrane-RNA binding protein, influenza virus matrix protein M1. Nat Struct Biol *4*, 239-244.

Snyder, P. M., Olson, D. R., McDonald, F. J., and Bucher, D. B. (2001). Multiple WW domains, but not the C2 domain, are required for inhibition of the epithelial Na+ channel by human Nedd4. J Biol Chem *276*, 28321-28326.

Springael, J. Y., and Andre, B. (1998). Nitrogen-regulated ubiquitination of the Gap1 permease of Saccharomyces cerevisiae. Mol Biol Cell *9*, 1253-1263.

Springael, J. Y., De Craene, J. O., and Andre, B. (1999). The yeast Npi1/Rsp5 ubiquitin ligase lacking its N-terminal C2 domain is competent for ubiquitination but not for subsequent endocytosis of the gap1 permease. Biochem Biophys Res Commun *257*, 561-566.

Staub, O., Dho, S., Henry, P., Correa, J., Ishikawa, T., McGlade, J., and Rotin, D. (1996). WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na+ channel deleted in Liddle's syndrome. Embo J *15*, 2371-2380.

Staub, O., Gautschi, I., Ishikawa, T., Breitschopf, K., Ciechanover, A., Schild, L., and Rotin, D. (1997). Regulation of stability and function of the epithelial Na+ channel (ENaC) by ubiquitination. Embo J *16*, 6325-6336.

Strack, B., Calistri, A., Accola, M. A., Palu, G., and Gottlinger, H. G. (2000). A role for ubiquitin ligase recruitment in retrovirus release. Proc Natl Acad Sci U S A *97*, 13063-13068.

Stricker, R., Mottet, G., and Roux, L. (1994). The Sendai virus matrix protein appears to be recruited in the cytoplasm by the viral nucleocapsid to function in viral assembly and budding. J Gen Virol *75*, 1031-1042.

Strous, G. J., and Govers, R. (1999). The ubiquitin-proteasome system and endocytosis. J Cell Sci *112*, 1417-1423.

Sudol, M. (1996). Structure and function of the WW domain. Prog Biophys Mol Biol *65*, 113-132.

Sudol, M., and Hunter, T. (2000). NeW wrinkles for an old domain. Cell *103*, 1001-1004.

Sullivan, N. J., Sanchez, A., Rollin, P. E., Yang, Z. Y., and Nabel, G. J. (2000). Development of a preventive vaccine for Ebola virus infection in primates. Nature *408*, 605-609.

Takada, A., Robison, C., Goto, H., Sanchez, A., Murti, K. G., Whitt, M. A., and Kawaoka, Y. (1997). A system for functional analysis of Ebola virus glycoprotein. Proc Natl Acad Sci U S A *94*, 14764-14769.

Takada, A., Watanabe, S., Ito, H., Okazaki, K., Kida, H., and Kawaoka, Y. (2000). Downregulation of beta1 integrins by Ebola virus glycoprotein: implication for virus entry. Virology *278*, 20-26.

Takada, A., and Kawaoka, Y. (2001). The pathogenesis of Ebola hemorrhagic fever. Trends Microbiol *9*, 506-511.

Timmins, J., Scianimanico, S., Schoehn, G., and Weissenhorn, W. (2001). Vesicular release of ebola virus matrix protein VP40. Virology *283*, 1-6.

VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M., and Wolberger, C. (2001). Molecular insights into polyubiquitin chain assembly: crystal structure of the Mms2/Ubc13 heterodimer. Cell *105*, 711-720.

Vanderzanden, L., Bray, M., Fuller, D., Roberts, T., Custer, D., Spik, K., Jahrling, P., Huggins, J., Schmaljohn, A., and Schmaljohn, C. (1998). DNA vaccines expressing either the GP or NP genes of Ebola virus protect mice from lethal challenge. Virology *246*, 134-144.

Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T., and Noel, J. P. (2000). Structural basis for phosphoserine-proline recognition by group IV WW domains. Nat Struct Biol *7*, 639-643.

VerPlank, L., Bouamr, F., LaGrassa, T. J., Agresta, B., Kikonyogo, A., Leis, J., and Carter, C. A. (2001). Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). Proc Natl Acad Sci U S A *98*, 7724-7729.

Vincent, S., Gerlier, D., and Manie, S. N. (2000). Measles virus assembly within membrane rafts. J Virol *74*, 9911-9915.

Voges, D., Zwickl, P., and Baumeister, W. (1999). The 26S proteasome: a molecular machine designed for controlled proteolysis. Annu Rev Biochem *68*, 1015-1068.

Volchkov, V. E., Becker, S., Volchkova, V. A., Ternovoj, V. A., Kotov, A. N., Netesov, S. V., and Klenk, H. D. (1995). GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. Virology *214*, 421-430.

Volchkov, V. E., Feldmann, H., Volchkova, V. A., and Klenk, H. D. (1998a). Processing of the Ebola virus glycoprotein by the proprotein convertase furin. Proc Natl Acad Sci U S A *95*, 5762-5767.

Volchkov, V. E., Volchkova, V. A., Slenczka, W., Klenk, H. D., and Feldmann, H. (1998b). Release of viral glycoproteins during Ebola virus infection. Virology *245*, 110-119.

Volchkova, V. A., Feldmann, H., Klenk, H. D., and Volchkov, V. E. (1998). The nonstructural small glycoprotein sGP of Ebola virus is secreted as an antiparallel-orientated homodimer. Virology *250*, 408-414.

Wang, G., Yang, J., and Huibregtse, J. M. (1999). Functional domains of the Rsp5 ubiquitin-protein ligase. Mol Cell Biol *19*, 342-352.

Wang, G., McCaffery, J. M., Wendland, B., Dupre, S., Haguenauer-Tsapis, R., and Huibregtse, J. M. (2001). Localization of the Rsp5p ubiquitin-protein ligase at multiple sites within the endocytic pathway. Mol Cell Biol *21*, 3564-3575.

Watanabe, S., Takada, A., Watanabe, T., Ito, H., Kida, H., and Kawaoka, Y. (2000). Functional importance of the coiled-coil of the Ebola virus glycoprotein. J Virol *74*, 10194-10201.

Weissenhorn, W., Carfi, A., Lee, K. H., Skehel, J. J., and Wiley, D. C. (1998). Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. Mol Cell 2, 605-616.

Will, C., Muhlberger, E., Linder, D., Slenczka, W., Klenk, H. D., and Feldmann, H. (1993). Marburg virus gene 4 encodes the virion membrane protein, a type I transmembrane glycoprotein. J Virol *67*, 1203-1210.

Wills, J. W., Cameron, C. E., Wilson, C. B., Xiang, Y., Bennett, R. P., and Leis, J. (1994). An assembly domain of the Rous sarcoma virus Gag protein required late in budding. J Virol *68*, 6605-6618.

Wilson, J. A., Bosio, C. M., and Hart, M. K. (2001a). Ebola virus: the search for vaccines and treatments. Cell Mol Life Sci *58*, 1826-1841.

Wilson, J. A., Bray, M., Bakken, R., and Hart, M. K. (2001b). Vaccine potential of Ebola virus VP24, VP30, VP35, and VP40 proteins. Virology *286*, 384-390.

Winberg, G., Matskova, L., Chen, F., Plant, P., Rotin, D., Gish, G., Ingham, R., Ernberg, I., and Pawson, T. (2000). Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3 protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases. Mol Cell Biol *20*, 8526-8535.

Xu, L., Sanchez, A., Yang, Z., Zaki, S. R., Nabel, E. G., Nichol, S. T., and Nabel, G. J. (1998). Immunization for Ebola virus infection. Nat Med *4*, 37-42.

Yang, Z., Delgado, R., Xu, L., Todd, R. F., Nabel, E. G., Sanchez, A., and Nabel, G. J. (1998). Distinct cellular interactions of secreted and transmembrane Ebola virus glycoproteins. Science *279*, 1034-1037.

Yang, Z. Y., Duckers, H. J., Sullivan, N. J., Sanchez, A., Nabel, E. G., and Nabel, G. J. (2000). Identification of the Ebola virus glycoprotein as the main viral determinant of vascular cell cytotoxicity and injury. Nat Med *6*, 886-889.

Yashiroda, H., Oguchi, T., Yasuda, Y., Toh, E. A., and Kikuchi, Y. (1996). Bul1, a new protein that binds to the Rsp5 ubiquitin ligase in Saccharomyces cerevisiae. Mol Cell Biol *16*, 3255-3263.

Yasuda, J., and Hunter, E. (1998). A proline-rich motif (PPPY) in the Gag polyprotein of Mason-Pfizer monkey virus plays a maturation-independent role in virion release. J Virol *72*, 4095-4103.

Zakowski, J. J., Petri, W. A., Jr., and Wagner, R. R. (1981). Role of matrix protein in assembling the membrane of vesicular stomatitis virus: reconstitution of matrix protein with negatively charged phospholipid vesicles. Biochemistry *20*, 3902-3907.

Zeller, H., and Bouloy, M. (2000). Infections by viruses of the families Bunyaviridae and Filoviridae. Rev Sci Tech *19*, 79-91.

Zhang, J., and Lamb, R. A. (1996). Characterization of the membrane association of the influenza virus matrix protein in living cells. Virology *225*, 255-266.

Zhang, J., Pekosz, A., and Lamb, R. A. (2000). Influenza virus assembly and lipid raft microdomains: a role for the cytoplasmic tails of the spike glycoproteins. J Virol *74*, 4634-4644.

Zhou, W., and Resh, M. D. (1996). Differential membrane binding of the human immunodeficiency virus type 1 matrix protein. J Virol *70*, 8540-8548.

Zoladek, T., Tobiasz, A., Vaduva, G., Boguta, M., Martin, N. C., and Hopper, A. K. (1997). MDP1, a Saccharomyces cerevisiae gene involved in mitochondrial/cytoplasmic protein distribution, is identical to the ubiquitin-protein ligase gene RSP5. Genetics *145*, 595-603.

Chapter 7: Materials and Methods

1. DNA cloning techniques

1.1 Preparation of DH5α competent cells

For cloning and amplification of DNA plasmids the bacterial strain, DH5α, is commonly used. The cells were originally purchased from Life Technologies, and were subsequently used to prepare fresh stocks of competent cells according to a protocol variant of one due to Mike Scott, Department of Neurology, UCSF. Starting from a single colony of DH5α, a 5 ml preculture of cells was grown overnight and used to inoculate 500 ml of TYM medium (2% Bacto-tryptone, 0.5% yeast extract, 0.1 M NaCl and 10 mM $MgSO_4$). The cells were grown at 37°C with rigorous shaking. When the cells reached an optical density (OD_{600}) of 0.6, the flask was put in ice water for rapid cooling of the cells, before centrifuging the cells at 4000 rpm for 10 minutes (min). The cell pellet was resuspended in 100 ml of TfBI (see below) by gently shaking on ice, and centrifuged for a further 10 min at 4000 rpm. Finally the cells were resuspended in 20 ml of TfBII (see below) before being aliquoted into pre-cooled microcentrifuge tubes.

TfBI: 30 mM KCH_3CO_2, 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$, 15% glycerol

TfBII: 10 mM Na-MOPS pH 7.0, 75 mM $CaCl_2$, 10 mM KCl, 15% glycerol

1.2 DNA preparation from DH5α cells

For the purpose of small-scale DNA plasmid purification, single colonies were grown overnight at 37°C in 3 ml of LB (1% Bacto-tryptone, 0.5% yeast extract, 1% NaCl, pH adjusted to 7.0 with 1 M NaOH). The pelleted cells were first resuspended in 200 µl of Buffer 1 (50 mM Tris-HCl pH 8.0, 10 mM ethylenediaminetetraacetic acid (EDTA), 100 µg/ml RNase A), before being lysed by the addition of 200 µl of Buffer 2 (200 mM NaOH, 1% SDS). After 5 minutes, the lysed cells were neutralised by adding 200 µl of Buffer 3 (3.0 M potassium acetate pH 5.5). The cell debris and genomic DNA were then pelleted by centrifuging the microcentrifuge tubes for 10 min at 13,000 rpm in a tabletop centrifuge. The protein contained in the supernatant was then removed by adding 1 volume of phenol/chloroform. After a further 5 min centrifugation, the upper aqueous phase was collected and the plasmid DNA was precipitated by the addition of 0.7 volumes of isopropanol. The precipitated DNA was pelleted by a 10 min centrifugation and washed with 1 ml of 70% ethanol. The plasmid DNA pellet was finally air-dried and resuspended in 50 µl of H_2O. For large-scale DNA purification the Qiagen plasmid Midi Kit was used, which is based on a modified alkaline lysis procedure, followed by purification of the plasmid DNA on Qiagen anion-exchange resin. Using this kit, the final DNA pellet was resuspended in 150 µl of H_2O, at a concentration of approximately 1 µg/µl.

1.3 Library amplification of cDNAs

When the cDNA of the gene of interest was not available in the laboratory, the cDNA was amplified from a cDNA library. This was the case for human clones of ADAM 19, Nedd4 and Tsg101. In the case of ADAM 19, the full human

sequence had not yet been deposited in the EMBL or GENBANK databases. Therefore, most of the human sequence was obtained by BLAST® search of the high throughput genomic sequences (htgs database) using the mouse ADAM 19 sequence. The human sequence retrieved from this search was used to design oligonucleotides to amplify human ADAM 19 from a cDNA library. The cDNA was amplified as two fragments of 1200 bp (5′ end) and 1900 bp (3′ end) respectively. A common restriction site (BamHI) was included in the reverse primer of the 5′end and in the forward primer of the 3′ end, so as to ligate the two pieces together. These fragments were amplified by polymerase chain reaction (PCR) using a human skeletal muscle cDNA library (Clontech) as a template. A typical 50 µl reaction included:

1:20 dilution of a human amplified skeletal muscle cDNA

20 pmol 5′ primer

20 pmol 3′ primer

5 µl 2 mM dNTPs

5 µl 10 x Pfu buffer

1 µl Cloned Pfu (Stratagene)

+ H_2O to a final volume of 50 µl.

The reactions were run in a Biometra thermocycler with the following programme: 5 min at 95°C, 35 cycles of 20 seconds at 95°C, 1 min at 65°C, 4 to 6 min at 72°C, and finally an extra 10 min at 72°C. Because the yield was not very good, the fragments were initially cloned using the Zero Blunt TOPO PCR cloning Kit (Invitrogen) into the pCR-BluntII-TOPO vector. Sequencing of the fragments revealed that there was a 400 bp deletion in the 5′ end. As a result it was decided to amplify only the disintegrin and cysteine-rich domains of ADAM 19 (878 bp)

from a different cDNA library, a human pancreas cDNA library (Clontech). The oligonucleotides were designed for cloning of ADAM 19 into pMelBac (Invitrogen) with a *myc*-epitope included in the 3′ primer. To obtain sufficient PCR product, a touch-down PCR was set up as follows:

5 min 95°C

5 cycles of: 30 seconds at 95°C, 1 min at 65°C and 3 min at 72°C

2 cycles of: 30 seconds at 95°C, 1 min at 63°C and 3 min at 72°C

30 cycles of: 30 seconds at 95°C, 1 min at 60°C and 3 min at 72°C

10 min at 72°C

The PCR fragment was subsequently cloned into the pMelBac vector by standard subcloning techniques, as described below.

In the case of Nedd4, a partial cDNA of human Nedd4 (NED4_HUMAN; accession number: D42055), lacking its N-terminal C2 domain (ΔC2) was amplified as two separate fragments, the first corresponding to the four WW domains and the second corresponding to the full HECT domain, from the human skeletal muscle cDNA library (Clontech) described above. A common restriction site (an endogenous XhoI site) was used to ligate the two fragments together, and the resulting product was used as a PCR template to amplify all four WW domains and the HECT domain. The 5′primer included an NheI site, while the 3′ primer included a HindIII site for cloning into pMal-c2g (New England Biolabs). The commercial pMalc2g vector used was modified so as to include a Tobacco etch virus (TEV) protease cleavage site at the 5′ end of the multiple cloning site (figure 66). The further Nedd4 deletion constructs were amplified by PCR using the ΔC2 construct as a template.

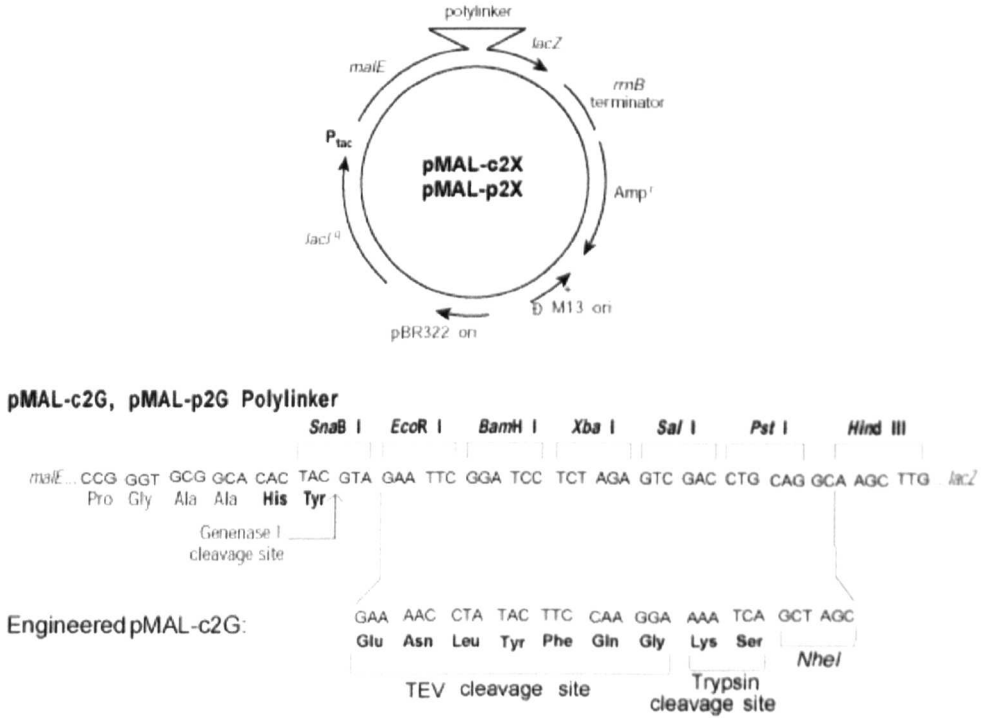

Figure 66: Vector map and multiple cloning site of the original and modified pMAL-c2g vector.

For Tsg101 (accession number: U82130) a full-length clone (1170 base pairs) was directly amplified from the human skeletal muscle cDNA library by standard PCR methods and was cloned into the NheI and HindIII sites of the modified pMalc2g vector. Subsequently, a construct corresponding to the N-terminal domain (residues 1-145) was also cloned into pMalc2g for expression as an MBP-fusion protein.

1.4 Polymerase chain reaction (PCR)

For a vast majority of the clones prepared in this work, standard subcloning procedures were followed. The template for the ADAM 12 clones was a full-length cDNA cloned into pBluescript (gift from U. Wewer). Similarly, a full-length cDNA clone of VP40 was used as a template for all the further VP40 constructs. The primers designed to amplify the constructs of interest were around 30 bp long, and had an average melting temperature of 65-70°C. They also included short

overhangs corresponding to the restriction sites and either an ATG (start codon) in the case of 5′ primers, or a stop codon in the case of 3′ primers. The primers were designed to obtain a PCR product with different restriction sites at either end. A standard PCR reaction included the following:

0.1–1 μg template DNA

10 pmol Primer 1

10 pmol Primer 2

5 μl 2mM dNTPs

5 μl 10 x Pfu Buffer

1 μl Cloned Pfu

+ H_2O to a final volume of 50 μl.

The cloned Pfu (Stratagene) polymerase was preferentially used for its proofreading capacity, resulting in PCR products with fewer errors than those obtained with the Taq polymerase. All the PCRs were run in a Biometra thermocycler with heated lids (no need to add oil). A typical PCR programme included an initial 5 min at 95°C, to fully denature the template DNA, followed by 25-30 cycles of 1 min at 95°C (denaturing step), 1 min at 55-65°C (annealing step; the temperature is determined by the length and melting temperatures of the oligonucleotides), and 1-4 min at 72°C (DNA synthesis step; the time depends on the length of the PCR product to be amplified; commonly for the Pfu polymerase one assumes 2 min/kilobase), and the programme usually included a final extension step at 72°C for 5 to 10 min.

The PCR reactions were then analysed by agarose gel electrophoresis. The larger the DNA fragment, the lower is the percentage of agarose needed to separate the DNA. Typically for DNA products ranging from 0.5 to 2 kilobasepairs

(kbp), a 1% (w/v) agarose gel is suitable. These gels were prepared by heating the weighed agarose (USB) in the appropriate volume of TAE (Tris Acetate EDTA) buffer (see below) to dissolve the powder. The mix was allowed to cool to approximately 60°C, before adding 10 μg of ethidium bromide. The gel mix was then poured into its gel cast (Pharmacia). Before loading the DNA on the gel, DNA loading buffer (see below) was added to the PCR reaction. DNA ladders (1kbp ladder from New England Biolabs) were usually run alongside the DNA of interest. The gel was commonly run in TAE buffer at 70-150 V, depending on the size of the gel.

TAE (50x): 50 mM EDTA, 5.7% (v/v) glacial acetic acid, 2 M Tris-base.

DNA loading buffer (10x): 0.1 M EDTA, 0.5% SDS, 0.1% (v/v) bromophenol blue, 25% glycerol.

The DNA was then visualised under ultra-violet (UV) light. To purify the PCR products from the gel, the Qiagen DNA Gel Extraction Kit was used and the DNA was eluted in 30 μl of H_2O.

1.5 Digestion of the DNA fragments and ligation

The PCR products, in addition to the corresponding plasmid DNA in which the PCR fragments were to be ligated, were then digested with the appropriate restriction enzymes (New England Biolabs) at 37°C for 2 hours. After the digestion, the plasmid DNA was purified on a 0.7% agarose gel as described above, and the PCR products were purified with the Qiagen Nucleotide Removal Kit. The amount of plasmid DNA and PCR product was then estimated by agarose gel electrophoresis alongside known amounts of DNA markers. Finally, the ligation reaction using the T4 ligase (New England Biolabs) was set up with a ratio

of insert to plasmid of 3:1. A control ligation including only the digested vector was also set up. The ligation reactions were incubated at 16°C for at least 4 hours.

1.6 Transformation of DH5α cells

The ligation reactions were then used for transformation of competent DH5α cells. For each transformation, 100 µl of cells were incubated on ice with the ligation reaction (10 µl) for 30 min, before the 5 min heat shock at 37°C. The cells were then grown for 1 hour in a shaker at 37°C in 900 µl of SOC medium (see below). All the cells were plated on LB-agar (15 g/litre of LB) plates with the appropriate selective antibiotic, to allow for growth of only recombinant clones. The plates were incubated for at least 12 hours at 37°C, until separate colonies could be picked and tested for the presence of the gene of interest. This was done by preparing small-scale DNA preparations of individual colonies (as described above) and digesting them with the appropriate restriction enzymes to determine whether they contained the desired insert or not. If the transformants were positive, large-scale DNA preparations were prepared, the correctness of the DNA sequence was verified by DNA sequencing (Genome Express) and the new clones could be tested for expression.

SOC medium: 2% Bacto-tryptone, 0.55% yeast extract, 10 mM NaCl, 10 mM KCl, and after autoclaving add 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose.

2. Expression systems

2.1 Expression in *E. coli*

2.1.1 Expression strains

For the expression in bacteria, two strains of *E. coli* were used in this work: BL21 (DE3) pUBS (Brinkmann et al., 1989) and BL21 (DE3) Codon Plus™ (Stratagene). Codon bias is a significant obstacle for efficient expression of heterologous genes in *E. coli* hosts. This problem has been most thoroughly documented for the arginine codons AGA and AGG, which are very rare codons in *E. coli*. The BL21 (DE3) pUBS strain contains a plasmid that encodes extra copies of the argU transfer RNAs (tRNA) gene and is able to rescue expression of genes restricted by the AGA or AGG codons. However, codons for isoleucine (AUA), leucine (CUA), and proline (CCC) are also known to affect the amount and quality of protein produced in *E. coli* hosts. The BL21 (DE3) Codon Plus strain contains a ColE1-compatible plasmid that encodes extra copies of the argU, ileY and leuW tRNA genes and is able to rescue expression of genes restricted by AGG/ AGA, AUA, and CUA codons respectively.

2.1.2 Expression tests

After transformation of these BL21 strains with the expression clones, single colonies were used for small-scale expression tests. A 2 ml culture of LB containing the appropriate antibody was inoculated with a single colony and grown to an optical density at 600 nm (OD_{600}) of 0.6. At this stage 1 ml of culture was removed and the left over culture was induced to express the recombinant protein by addition of 1 mM isopropyl-b-D-thiogalactopyranoside (IPTG; Eurogentec). The culture was grown in a shaker (180 rpm) at 37°C for 2 to 3 hours.

150 µl of uninduced and induced cells were then pelleted and solubilised in SDS-loading buffer for SDS-PAGE analysis. The left over uninduced cells (850µl) were mixed with 150 µl of glycerol and frozen in dry ice and ethanol to be stored at -80°C.

2.1.3 Large-scale cultures

For large-scale cultures, a 2 ml preculture was first prepared. Scraped cells from the glycerol stocks were directly used to inoculate small volumes of LB in 14 ml Falcon tubes. Once the cells had reached the exponential phase, they could be used to inoculate large flasks of LB, by adding 0.5-1 ml of the preculture to 1 litre of autoclaved LB containing the appropriate antibiotics. The culture was left over night at room temperature and without shaking. The following day, the cultures were grown in a shaker (180 rpm) at 37°C and induced with IPTG at an OD_{600} of 0.6. In some cases, the low solubility of the recombinant protein required that the induction be at a lower temperature (e.g. 30°C).

After a 3-hour induction, the cells were pelleted by centrifugation for 20 min at 5000 rpm, before being resuspended in an appropriate buffer, the content and volume of which varies according to the properties (such as its pI) of the recombinant protein. A cocktail of protease inhibitors (Roche) and 100 mM phenylmethylsulfonyl fluoride (PMSF) were usually added to the cell suspension prior to the lysis. The cells were then lysed in a metallic beaker, on ice, by 5-10 cycles of 30 seconds of sonication with a Misonix ultrasonic processor XL. The crude extract was then cleared by centrifugation at 20,000 rpm for 45 min. The supernatant obtained after centrifugation corresponded to the soluble protein

fraction, which could be used for further chromatography steps. In the case of insoluble protein, the pellet could be used for purification of inclusion bodies.

The following antibiotic stocks (BIO101) were used at a 1:1000 dilution:

Ampicillin: 100 mg/ml

Kanamycin: 25 mg/ml

Chloramphenicol: 34 mg/ml

2.2 Expression in *Pichia pastoris*

2.2.1 *Pichia pastoris* strains and storage

The following strains of *Pichia pastoris* (Invitrogen) were used in this study:

Strain	Genotype	Phenotype
X-33	wild-type	Mut+
GS115	*his4*	His-, Mut+
KM71H	*arg4 aox1::ARG4*	Mut^S, Arg+

Pichia pastoris is a methylotrophic yeast, capable of metabolising methanol as its sole carbon source. The first step in the metabolism of methanol is the oxidation of methanol to formaldehyde using molecular oxygen by the enzyme alcohol oxidase. The promoter regulating the production of alcohol oxidase (AOX1 gene) is the one used to drive heterologous protein expression in *Pichia*. Expression of the recombinant protein is induced by methanol. For long term storage, the cells were grown overnight in YPD medium (see below), harvested and resuspended in YPD containing 15% glycerol at a final OD_{600} of 50-100, before being frozen in dry ice and ethanol for storage at -80°C. All of the work with *Pichia* was essentially done according to the EasySelect *Pichia* Expression Kit manual (Invitrogen). The cells were grown either on agar plates or in liquid medium in baffled flasks covered with 2-3 layers of sterile cheesecloth layers at 30°C.

2.2.2 Transformation of *Pichia pastoris*

Prior to transformation and selection in *Pichia*, the plasmid DNA (pPICZα) containing the gene of interest needed to be linearised by digestion with SacI. The DNA was then purified and ethanol precipitated using 1/10 volume 3 M sodium acetate and 2.5 volumes of 100% ethanol. After washing the pellet with 80% ethanol, the DNA was resuspended in sterile H_2O. To prepare the cells for transformation, 500 ml of fresh YPD medium was inoculated with 0.1-0.5 ml of an overnight culture of each *Pichia* strain and grown for another 24 hours to reach an OD_{600} of 1.3-1.5. The cells were then harvested, washed twice with 250 ml ice-cold water, once with 20 ml of ice-cold 1 M sorbitol, before being resuspended in 1 ml of 1 M sorbitol. 80 µl of these cells were then mixed with 10 µg of linearised DNA and transferred to a 0.2 cm electroporation cuvette to be pulsed for 10 ms with a field strength of 1500 V. 1 ml of 1 M sorbitol was immediately added to the cells. After 1-2 hour incubation at 30°C, without shaking, various amounts of cells (10, 25, 50, 100 and 200 µl) were spread on YPDS plates (see below) containing 100 µg/ml Zeocin. An extra 200 µl of cells were spread on YPDS plates containing increasing amounts of Zeocin (500, 1000 and 2000 µg/ml). Recombinant colonies appeared after 2-4 days.

2.2.3 Determination of phenotype and expression test

To determine whether the X-33 and GS115 transformants were Mut+ or Mut^S, single colonies were picked and streaked in a regular pattern on both an MMH plate and an MDH plate (see below). Mut+ strains will grow normally on both plates, while Mut^S strains will grow normally on the MDH plate but show little or no growth on the MMH plate. Expression of the gene of interest was then

tested in the GS115 strain. For this purpose, 25 ml of MGYH and BMGH medium (see below) was inoculated with single zeocin-resistant colonies and grown in a shaker (290 rpm) at 30°C until the OD_{600} reached 2-6. The cells grown in MGYH medium were then diluted in MMH medium to an OD_{600} of 1. Similarly, the cells grown in BMGH medium were transferred to BMMH medium. The transfer to methanol-containing medium induces the expression of the gene of interest. 100% methanol was added to a final concentration of 0.5% every 24 hours to maintain induction. Every 12 hours, 1 ml of the expression culture was removed to analyse the expression level in the supernatant and the cells. To analyse the cell pellets, the cells from each 1 ml aliquot were resuspended in 100 µl of breaking buffer (see below). An equal volume of acid-washed glass beads (size 0.5 mm) was added and the cells were lysed by 8 cycles of 30 seconds of vortexing. After centrifuging the cell extract at maximum speed in a tabletop centrifuge for 10 min at 4°C the cleared supernatant was collected and could be used for protein detection.

Pichia media and buffers:

YPD (Yeast Extract Peptone Dextrose Medium): 1% yeast extract, 2% peptone, 2% dextrose (glucose)

YPDS plates: YPD + 1 M sorbitol, 2% agar

MGYH (Minimal Glycerol Medium with Histidine): 1.34% yeast nitrogen base (YNB), 1% glycerol, $4x10^{-5}$% biotin, 0.004% histidine

MDH (Minimal Dextrose Medium with Histidine): 1.34% YNB, $4x10^{-5}$% biotin, 2% dextrose, 0.004% histidine

MMH (Minimal Methanol with Histidine): 1.34% YNB, $4x10^{-5}$% biotin, 0.5% methanol, 0.004% histidine

BMGH (Buffered Minimal Glycerol with Histidine): 100 mM potassium phosphate pH 6.0, 1.34% YNB, $4x10^{-5}$% biotin, 1% glycerol

BMMH (Buffered Minimal Methanol with Histidine): 100 mM potassium phosphate pH 6.0, 1.34% YNB, $4x10^{-5}$% biotin, 0.5% methanol

Breaking buffer: 50 mM sodium phosphate, pH 7.4, 1 mM PMSF, 1 mM EDTA, 5% glycerol

2.3 Expression in insect cells

2.3.1 Maintenance of Sf9 cells

The Sf9 (ovarian cell derived *Spodoptera frugiperda*) insect cells were cultured at 28°C and grown either as monolayers in plastic cell culture flasks (75 cm^2) or in suspension in spinner flasks (could contain up to 1 litre of cell suspension). The cells were originally grown in TNM-FH (Sigma) medium supplemented with 10% foetal calf serum, but were progressively transferred to serum-free medium (SF900II; Life Technologies) containing 50 units/ml penicillin and 50 μg/ml streptomycin (Sigma). A confluent cell layer corresponds to a cell density of $2x10^6$ cells/ml. The cell density was determined using a Neubauer counting chamber. The cells were diluted 1:1 in trypan blue (Sigma) and 10 μl of the mix was used to count the number of cells within 16 squares of the hemacytometer. This number is converted to a cell density (cells/ml) by multiplying it by $2x10^4$. The cells were subcultured at confluence by dislodging the cells and diluting them 5-10 times in fresh medium. In spinner flasks, when the cells reached a density of $2x10^6$ cells/ml fresh medium was added to the flask to dilute the cells; if needed the cell suspension was transferred to a larger spinner flask.

2.3.2 Transfection of Sf9 cells

For expression in Sf9 cells the protocols described in the Bac-N-Blue Transfection Kit manual (Invitrogen) were followed. For transfection of the cells, $0.8x10^6$ cells (with greater than 98% viability) were seeded in 6-well plates and left to attach. 4 μg of the recombinant transfer plasmid (pMelBac) containing the gene of interest, 1 ml of Grace's insect medium and 20 μl of Insectin-Plus™ liposomes were added to a microcentrifuge tube containing 0.5 μg of Bac-N-Blue™ DNA and the transfection mixture was vortexed vigorously for 10 seconds followed by a 15 min incubation at room temperature. Meanwhile the cells were washed with Grace's insect medium, before the addition drop wise of the transfection mixture (one transfection mixture was used for two wells containing each $0.8x10^6$ cells). The plates were then incubated for 4 hours at room temperature, before addition of 0.5 ml of fresh culture medium to each well. After a 72-hour incubation at 28°C, the medium was collected (transfection viral stock) and used to isolate recombinant baculoviruses.

2.3.3 Plaque assay

The same method, namely the plaque assay, was followed to isolate recombinant baculovirus and to titre new viral stocks. For the plaque assay $5x10^6$ cells were seeded per 100 mm plate at 50% confluence. Once the cells had attached, various dilutions (10^{-2}, 10^{-3}, 10^{-4}, 10^{-6}, 10^{-8}) of the viral stocks were used to infect the cells for 1 hour at room temperature. The medium was then removed and the cells were overlaid with a medium/agarose mix to "pin down" the infected monolayer. The medium/agarose mix consisted of a 1:1 mix of culture medium containing 5-bromo-4-chloro-3-indolyl-b-D-galactoside (X-gal; Life

Technologies) at a final concentration of 150 μg/ml, and an agarose-medium mix (1:1 mix of 2.5% agarose (Invitrogen) and fresh medium) maintained at 47°C. Once the agarose had set, the plates were sealed and incubated at 28°C until clear blue plaques (corresponding to recombinant baculovirus) could be isolated. To determine a virus titre from a plaque assay, the plaques were counted on a plate that had between 50-100 plaques and the virus titre (pfu/ml) was calculated as: (1/dilution) x number of plaques. A good virus stock has a titre of 1-2x10^8 pfu/ml.

2.3.4 Virus stock amplification

To prepare an original small-scale high-titre P1 virus stock, single blue plaques obtained from the plaque assay following transfection, were added to 5x10^6 cells seeded in 6-well plates. The agarose plugs containing the plaques were transferred to the wells using sterile Pasteur pipettes. The virus stock was then collected after several days at 28°C, when all the cells had lysed (5-8 days). After titration (as described above) this P1 virus stock was amplified to obtain a large-scale high-titre virus stock (1 litre). For this purpose, Sf9 cells were amplified in spinner flasks so as to obtain a 1-litre suspension at a cell density of 1.5x10^6 cells/ml. These cells were infected with the P1 virus stock at a multiplicity of infection (MOI) of 0.5. The medium was then collected after 7-10 days, once over 90% cells had lysed. For expression purposes, however, confluent cells (either as monolayers or in suspension) were infected at a MOI of 5 and incubated at 28°C for 72 hours before harvesting the cells and supernatant.

3. Protein purification

3.1 Inclusion body purification

When expressed in *E. coli*, many recombinant proteins are insoluble and are present as inclusion bodies. These insoluble aggregates of misfolded protein can, nevertheless, be purified from the pellet obtained after the centrifugation at 20,000 rpm for 45 min (see section 2.1). This was the case for the purification of the bacterially expressed ADAM 12 constructs. The pellets resulting from a 3-litre culture were resuspended mechanically (on a rotator) in 100 ml of washing buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5% Triton X100, 1 mM EDTA and 5 mM DTT) for 1 hour at room temperature. After centrifuging the suspension at 20,000 rpm for 25 min at 4°C, the washes were repeated 2-3 times. Finally the pellet was washed with 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA and 5 mM DTT to remove the triton, and resuspended in 15 ml of 25 mM MES pH 6.0, 8 M guanidine-HCl and 5 mM DTT. The protein could then be stored at -80°C.

3.2 Refolding

To obtain soluble protein, two refolding strategies were tested. In the first method the denatured protein was diluted in 100 ml of 50 mM Tris-HCl pH 7.5, 6 M urea, 5 mM reduced glutathione and 0.5 mM oxidised glutathione, to a final protein concentration of 2 μM. The presence of the reduced and oxidised glutathione allows for the formation of disulphide bonds, which cannot form within the bacteria. After a 24-hour incubation of this mix at 4°C, the protein solution was centrifuged at 8000 rpm for 10 min and the supernatant was dialysed against a large volume (5 litres) of 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 2 M urea overnight at 4°C. Finally, to remove any residual urea, the protein solution

was further dialysed against 50 mM Tris-HCl pH 7.5 and 100 mM NaCl. The second method is known as 'refolding by dilution'. The protein was directly added drop wise to a stirring solution of 100 ml of 50 mM Tris-HCl pH 7.5, 100 mM NaCl, 400 mM L-Arginine-HCl, 5 mM reduced glutathione and 0.5 mM oxidised glutathione, to a final protein concentration of 2 μM. After 24 hours of refolding at 4°C, the protein was dialysed against 5 litres of 25 mM Tris-HCl pH 7.5 and 150 mM NaCl. In both cases, the aggregated protein was pelleted by centrifugation, and the protein in the supernatant could be quantified and used for further purification steps.

3.3 Expression and purification of hADAM 12 from Sf9 cells

A general description of the protocol used for the purification of the disintegrin and cysteine-rich domains of ADAM 12 is given in Chapter 3 section 2.1. After the ammonium sulfate precipitation, ADAM 12, which has a C-terminal His-tag, was further purified on an immobilised metal ion affinity column. For this purpose, 10 ml of chelating sepharose fast flow (Pharmacia) gel was packed into a column (Biorad) and loaded with 2 column volumes of 100 mM $NiSO_4$. After washing the gel with 5 column volumes of distilled H_2O, the column was equilibrated with 3 column volumes of 50 mM Tris-HCl pH 8.0 and 100 mM NaCl. ADAM 12 was then loaded on the column overnight at 4°C at 0.5 ml/min, using a BioLogic low-pressure chromatography system (Biorad). The next day, the column was washed at 1.5 ml/min with 5 column volumes of equilibration buffer, 2 column volumes of 5% Buffer 1 (see below) and finally with 10 column volumes of 10% Buffer 1. The protein was then eluted with an imidazole gradient from 15 to 100% of Buffer 2 (see below) and collected in 2 ml fractions. The resin was then

regenerated by successive washes with 50 mM EDTA, 0.5 M NaOH and H_2O, before applying fresh $NiSO_4$.

Buffer 1: 50 mM Tris-HCl pH 8.0, 100 mM NaCl and 270 mM Imidazole

Buffer 2: 50 mM Tris-HCl pH 8.0, 100 mM NaCl and 1 M Imidazole

After SDS-PAGE analysis, the fractions containing ADAM 12 were pooled and concentrated in 15 ml Centripreps and 2 ml Centricons (Amicon) for a final size exclusion chromatography step. 0.5-1 ml of concentrated ADAM 12 (2-3 mg) was separated by Fast Performance Liquid Chromatography (FPLC) on a pre-equilibrated Superdex 75 (Pharmacia) column at 0.5 ml/min in 50 mM Tris-HCl pH 8.0 and 100 mM NaCl.

3.4 Expression and purification of hNedd4 and hTsg101

The N-terminal construct of human Tsg101 and all the human Nedd4 constructs were expressed as fusion proteins with the maltose-binding protein (MBP). Tsg101 was expressed in *E. coli* strain BL21 (DE3) pUBS (Brinkmann et al., 1989), while the Nedd4 constructs were expressed in *E.coli* strain BL21 (DE3) Codon+ (Stratagene). Expression of the MBP fusion proteins was induced with 1 mM (IPTG) for 3 hours at 37°C. Cells were lysed by sonication in 50 mM Bicine pH 9.5, 100 mM NaCl (plus 10 mM DTT in the case of Tsg101 and the Nedd4 constructs containing the HECT domain). The supernatants were cleared by centrifugation before being loaded on a pre-equilibrated amylose resin (New England Biolabs), which binds the MBP. The binding capacity of the amylose resin is 3 mg/ml. After washing the column with 5-10 column volumes of loading buffer, the fusion proteins were finally eluted with the loading buffer containing 10 mM maltose. In the case of Nedd4 WW234-HECT, a large proportion of the

fusion protein was present in the amylose column wash, and comparison of the protein present in the wash with the protein in the elutions revealed that the protein in the wash was more soluble and presumably properly folded. This protein fraction was, therefore, used for the following chromatography step.

To remove the fusion protein, the eluted protein was concentrated and incubated overnight at 4°C with the TEV protease at a 1:100 (w/w) ratio. hNedd4 and Tsg101 could then be separated from the MBP by size exclusion chromatography on a pre-equilibrated Superdex 200 (Pharmacia) column. hNedd4 WW234-HECT was purified on a Q-Sepharose (Pharmacia) resin and eluted with a 0.1 M to 0.7 M NaCl gradient, to remove the MBP, prior to the size exclusion chromatography step.

3.5 Expression and purification of VP40

All VP40 constructs were cloned into the pRSET (NdeI/HindIII sites; Invitrogen) vector for expression in *E.coli* strain BL21 (DE3) pUBS (Brinkmann et al., 1989). The expression and purification of VP40 (31-326) was as described previously (Dessen et al., 2000). VP40 (1-326) expression was induced with 1 mM IPTG for 4 hours at 30°C. Cells were lysed by sonication in a buffer containing 50 mM Tris pH 8.8, 100 mM NaCl and 10 mM DTT, and the supernatant was cleared by centrifugation. VP40 (1-326) was then loaded on a Q-sepharose resin (Pharmacia), and the flow-through was collected, which contained the protein. Subsequently, VP40 (1-326) was pelleted by the addition of 50% (w/w) ammonium sulfate in a 2:1 ($V_{protein}/V_{(NH4)2SO4}$) ratio. After 1 hour at 4°C, the precipitated protein was collected by centrifugation and dissolved in 50 mM

Bicine pH 9.3, 100 mM NaCl and 10 mM DTT, before further purification by size exclusion chromatography (Superdex 200; Pharmacia).

Both VP40 (1-212) and VP40 (1-319) were copurified with the Nedd4(WW234)-MBP fusion protein. The expression of both VP40s was induced with 1 mM IPTG for 3 hours at 37°C. The cell pellets of the VP40 expressing bacteria and those of Nedd4(WW234)-MBP expressing bacteria were resuspended and sonicated together in 50 mM Bicine pH 9.5, 100 mM NaCl and 10 mM DTT. The supernatants were cleared by centrifugation and loaded on an amylose resin. After extensive washing of the resin, the proteins were coeluted with 10 mM maltose. The proteins were then concentrated and separated by size exclusion chromatography (superdex 200; Pharmacia) in the lysis buffer.

For the surface plasmon resonance studies, VP40 (1-212) and VP40 (1-326) were cloned into the pET21d vector for expression in BL21 (DE3) pUBS with a C-terminal His-tag. The expression of the His-tagged VP40 constructs was induced with 1 mM IPTG for 3 hours at 37°C. The cell pellets were resuspended in 50 mM Tris pH 8.8, 100 mM NaCl and 10 mM DTT (for VP40 (1-326)), and sonicated as described previously. The cleared supernatants were then loaded on 5 ml of chelated sepharose resin, loaded with 100 mM $NiSO_4$, washed with 20 ml of loading buffer and 20 ml of loading buffer containing 50 mM Imidazole. In the case of VP40 (1-212) the protein was found in the 50 mM Imidazole wash fraction. For VP40 (1-326), the protein was eluted with 500 mM Imidazole. The proteins were then concentrated and further purified by size exclusion chromatography on a Superdex 200 column (Pharmacia), in 50 mM Bicine pH 9.3, 100 mM NaCl and 10 mM DTT.

4. Protein characterisation

4.1 Polyacrylamide Gel Electrophoresis

Protein concentrations were estimated either by measuring the absorbance at 280 nm and the Lambert-Beer law $A = \varepsilon.l.C$ (where A is the absorption, ε the molar absorption coefficient ($M^{-1}.cm^{-1}$), l the pathlength (cm) and C the protein concentration (M)), or by using the Bradford assay (Biorad). In the latter, 1-20 µl of protein is added to 1 ml of pre-diluted Bradford solution (1:5 in distilled H_2O) and the absorbance at 595 nm is measured and compared to a protein standard series (usually made with bovine serum albumin, BSA).

At most stages during their purification, the proteins were analysed by denaturing (with sodium dodecyl sulfate or SDS) polyacrylamide gel electrophoresis (PAGE). 8 to 15% gels were prepared as described in (Sambrook J., 1989). Biorad mini-Protean II gel systems were used to run gels at 150-200 V. The following SDS-PAGE solutions were used:

Running buffer (10x) pH 8.3: 30 g/l Tris base, 144 g/l glycine and 10 g/l SDS

SDS sample buffer (4x) for 8 ml: 3.8 ml H_2O, 1 ml 0.5 M Tris-HCl pH 6.8, 0.8 ml glycerol, 1.6 ml 10% (w/v) SDS, 0.4 ml β-mercaptoethanol and 0.4 ml 1% (w/v) bromophenol blue.

When protein samples were run under non-reducing conditions, the β-mercaptoethanol was omitted from the sample buffer.

To separate proteins according to their charge, as opposed to their molecular weight in the case of SDS-PAGE, proteins were run on 8% native gels at pH 8.8, with no stacking gels. For non-denaturing PAGE, the gel mix, the running buffer and the sample buffer were prepared without SDS.

Native pH 8.8 sample buffer (5x): 0.3 M Tris-HCl pH 8.8, 50% glycerol and 0.05% (w/v) bromophenol blue.

To visualise the protein on polyacrylamide gels, the gels were stained with a coomassie staining solution (0.1% coomassie blue, 40% ethanol and 10% acetic acid) for 15 min, before incubating the gels in a destaining solution (5% acetic acid, 7% ethanol).

4.2 Western blotting

When the amounts of separated protein were not detectable by coomassie staining, Western blots (or immunoblots) were prepared. For this, SDS-PAGE gels were incubated in blotting buffer (see below) before transferring the protein to a nitrocellulose membrane (Schleicher & Schuell) using a Semiphor (Pharmacia) semi-dry transfer apparatus. The transfers were run at a constant current (between 80 and 150 mA depending on the size of the membrane) for 30-60 min. To ensure that all of the protein had been transferred, pre-stained protein molecular markers (Biorad) were loaded alongside the protein samples. Once the proteins were fully transferred to the membrane, the latter was incubated for 15 min in a blocking solution (see below). The membrane was then incubated with the primary antibodies diluted appropriately in blocking solution at room temperature for at least one hour. The membrane was then washed 3 times with PBT (see below) before being incubated for a further one hour with the secondary antibody (either anti-rabbit or anti-mouse) conjugated to alkaline phosphatase (Promega), diluted 1:7500 in PBT. The protein bands were then revealed after a further 3 washes with PBT and the addition of the Western Blue® stabilised alkaline phosphatase substrate (Promega). The following primary antibody dilutions were used:

- Monoclonal mouse anti-His (Qiagen): 1:2000 dilution in PBS + 3% BSA
- Polyclonal rabbit anti-ADAM 12 (gift from U. Wewer): 1:5000 dilution in 5% milk
- Monoclonal mouse anti-*myc* (Invitrogen): 1:5000 dilution in 5% milk
- Polyclonal rabbit anti-VP40 (gift from V. Volchkov): 1:1000 dilution in 5% milk

Blotting buffer: 48 mM Tris base, 39 mM glycine, 0.05% SDS and 20% methanol

Blocking solution: 5% non-fat dried milk in PBS

PBT: PBS and 0.02% Tween 20

PBS (phosphate buffered saline): 0.05 M phosphate buffer ($Na_2HPO_4 \cdot 7H_2O$ and $NaH_2PO_4 \cdot 1H_2O$), 0.137 M NaCl, pH 7.45

When the protein quantities analysed were too small for detection by western blot, the amplified Opti-4CN kit (Biorad) was used, which significantly amplifies the signal so as to detect as little as 1 or 2 ng of protein (equivalent to the sensitivity achieved with chemiluminescence).

4.3 Deglycosylation of ADAM 12

The deglycosylation enzymes, PNGase F and EndoH were purchased from New England Biolabs. The protein substrate for the deglycosylation enzymes was either native or denatured by heating the sample at 100°C for 10 min in denaturing buffer (0.5% SDS, 1% β-mercaptoethanol). Digestion of 0.5 µg of purified ADAM 12 protein with both PNGase F and EndoH was achieved by incubating the following reaction at 37°C for one hour:

- 5 µl of protein (native or denatured; 0.1 µg/µl)
- 3 µl of 10 x G7 buffer (0.5 M sodium phosphate pH 7.5)

- 3 µl of 10% NP40
- 1 µl EndoH (500 units) and 1 µl of PNGase F (20 units)

+ sterile H_2O to a final volume of 30 µl.

The reactions were then directly analysed on a 10% SDS-PAGE gel, followed by a western blot using the anti-ADAM 12 antibody.

4.4 Immunoprecipitation of ADAM 12 and ADAM 19

For the immunoprecipitation reaction, a protein solution (300 µl) containing both ADAM 12 and ADAM 19 (eluate from copurification on a Ni-affinity column) was used. In a first step, the protein solution was incubated with 2 µg of mouse monoclonal anti-*myc* antibody (Invitrogen) for 2 hours at room temperature on a rotator. A control tube missing the antibody was also prepared. Secondly, 20 µl of pre-washed and equilibrated protein A sepharose CL-4B beads (Pharmacia) were added to each tube for a further one hour incubation at room temperature. Finally the beads were washed 3 times with 1 ml of protein buffer (50 mM Tris-HCl pH 8.0 and 100 mM NaCl), and were finally resuspended in 60 µl of SDS sample buffer and boiled for 5 min to release the bound protein. The samples were loaded on two gels (30 µl/gel) for Western blot analysis with both the anti-*myc* and the anti-His antibodies.

4.5 Sucrose floatation and liposome binding assays

Samples for the floatation experiments were prepared by adding the sucrose directly to the collected cell supernatants to a final concentration of 50% sucrose (w/v). A discontinuous gradient was then overlaid on the 5 ml sucrose-containing sample. The gradient consisted of 2 ml of 40% sucrose (w/v), 2 ml of

35% sucrose (w/v), 2 ml of 20% sucrose (w/v) and finally 1 ml of 10% sucrose (w/v), all prepared in PBS. Centrifugation was performed in a SW 41 rotor at 40,000 rpm for 15 hours at 12°C. Fractions of 1 ml were collected from the gradients and 500 µl of these fractions were precipitated with 3.5% trichloroacetic acid (TCA). These samples were separated on a 12% SDS-PAGE and blotted onto nitrocellulose membranes for Western blot analysis. The VP40 bands were visualised using a VP40-specific anti-serum, as primary antibody and a secondary anti-rabbit-IgG conjugated to alkaline phosphatase.

For the liposome-binding assay, the supernatant of infected Sf9 cells expressing the disintegrin and cysteine-rich domains of ADAM 12 was concentrated ten times and incubated for 30 minutes at room temperature with 300 µl of liposomes (25% cholesterol, 25% phosphatidylcholine and 50% L-α-phosphatidyl-L-serine; lipid concentration at 2 mg/ml). The protein-liposome mix was then analysed on a discontinuous sucrose gradient as described previously (Ruigrok et al., 2000).

4.6 Crystallisation trials

For crystallisation trials, the protein corresponding to a single peak from the gel filtration chromatography step was concentrated to 8-10 mg/ml. Initial crystallisation trials were done with commercial screens (Hampton Research). 24-well plates (Hampton Research) were used for setting up the trials, using the hanging drop method with siliconised glass coverslips (Hampton Research). The protein was diluted 1:1 in the drop. Once interesting conditions had been found, the additive screens (Hampton Research) were also tested.

4.7 Circular dichroism

Spectral acquisition was performed on a Dichrograph Jobin Yvon CD6 spectropolarimeter, with a thermostated sample holder, between 190 and 260 nm. Purified ADAM 12 was placed in a linear polarised beam in a 0.1 cm quartz cuvette at a final concentration of 0.7 mg/ml in a phosphate buffer (PBS). For the measuring of the CD spectra at increasing temperatures, incubation times at each temperature were kept sufficiently long to reach the final state of denaturation. The mean residue ellipticity, θ(deg.cm2.dmol-1) was calculated according to:

$[\theta] = \frac{\theta_{obs} mrw}{10dc}$ where θ_{obs} is the observed dichroic absorption, *mrw* is the mean residue weight (110), *d* is the optical pathlength in centimetres, and *c* is the protein concentration in mg/ml.

5. Analysis of protein-protein interactions

5.1 "Pull-down" assay

To test the interaction of ADAM 12 with ADAM 19, the cell culture supernatants containing both of these proteins were precipitated with 55% ammonium sulfate. Each protein pellet was resuspended in 1 ml of 50 mM Tris-HCl and 100 mM NaCl. ADAM 12 was then loaded on a Ni-NTA spin column (Qiagen), washed with 3 column volumes of buffer, before ADAM 19 was applied to the column. Finally, after a further 8-10 column volumes of washing, the proteins were eluted with 0.5 M imidazole, 50 mM Tris-HCl and 100 mM NaCl. The eluate was further analysed by size exclusion chromatography (Superdex 200; Pharmacia), to see whether the proteins co-eluted.

Each purified VP40 construct was in turn mixed with either purified Nedd4(WW234)-MBP or Tsg101(N-term)-MBP, loaded on an amylose resin, washed extensively, and finally eluted with 10 mM maltose. The amount of protein in the elution fractions was quantitated with the Bradford assay (Biorad) and electrophoresed by SDS-PAGE on 12% gels.

The binding assays with the different hNedd4 constructs and VP40 (1-212) consisted of the mixing of the cell pellets of bacteria expressing VP40 (1-212) to each of the different hNedd4 construct-expressing bacteria independently. The bacteria were lysed together and the cleared supernatants were then loaded on the amylose resin as described above.

5.2 Surface plasmon resonance biosensor (BIAcore Upgrade)

Biosensors detect molecules and monitor binding events between two or more molecules, in real time, with high selectivity on the basis of molecular recognition. The central feature is a selective active surface consisting of a biological species coupled to an optically or electronically active medium. This technology relies on the phenomenon of surface plasmon resonance (SPR), which occurs when surface plasmon waves are excited at a metal/liquid interface (Karlsson and Falt, 1997). Surface plasmon resonance (figure 67B) occurs at a critical angle of incident light, and light energy is transferred to electrons in the metal film surface, causing a minimum in the reflected light. By measuring small changes in refractive index, the instrument monitors the change in mass as a ligand binds to, or dissociates from, its binding partner. In general, the refractive index change for a given change of mass concentration at the surface layer is

practically the same for all proteins and peptides. Figure 67A shows the instrumental set-up.

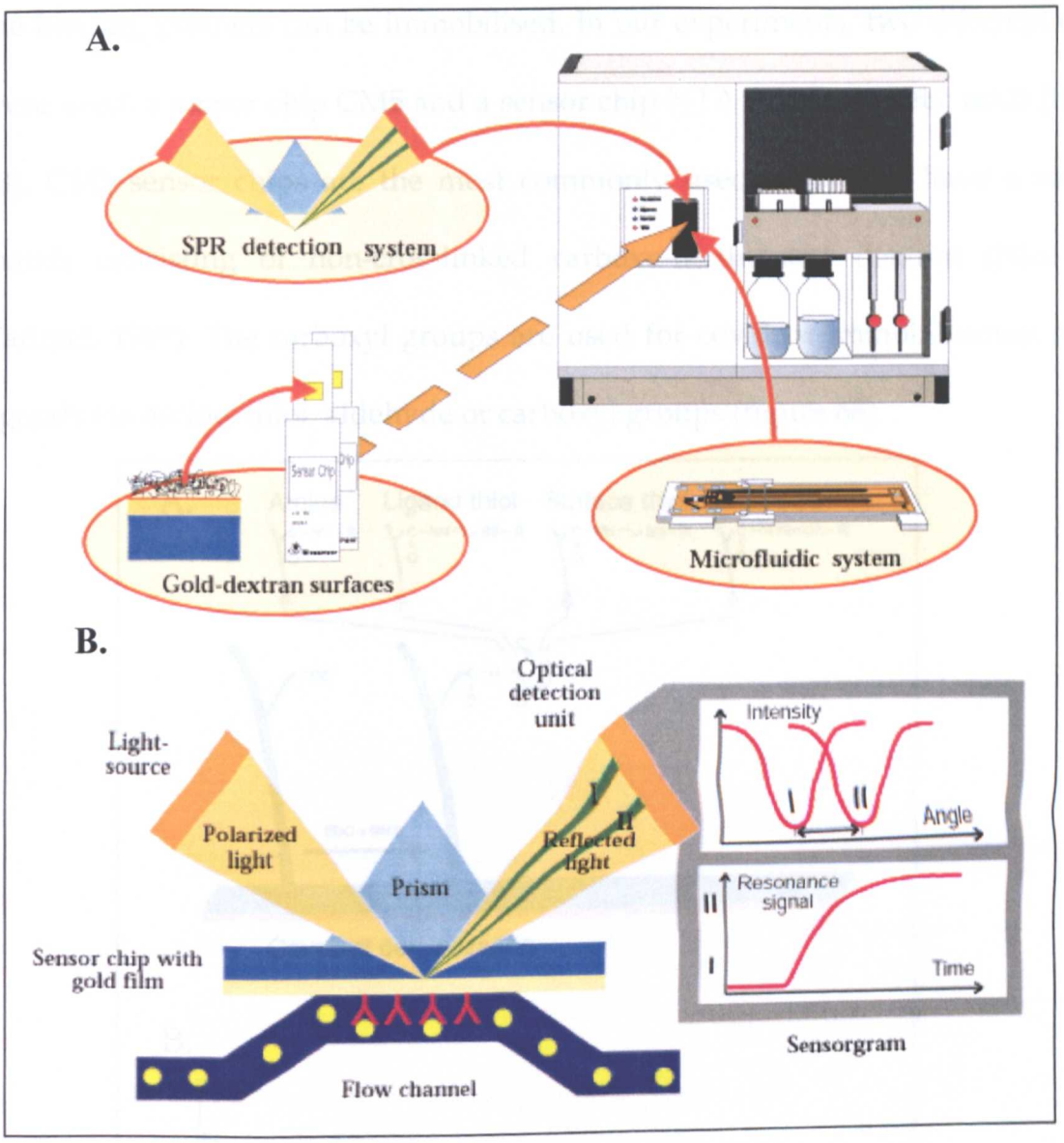

Figure 67: (A) Schematic representation of the instrumental set-up of the BIAcore 1000 biosensor. (B) Principle of surface plasmon resonance. The immobilised ligand is shown in red ( ) and the injected analyte in yellow ( ). Taken from http://www.biacore.com/biomol/pdfs.shtml.

Flow cells are formed by interfacing the sensor chip with a thermostatically controlled integrated fluidic cartridge. Four parallel channels (60 nl volume) are formed on the sensor surface. Each channel is used independently in the BIAcore 1000.

The sensor chip consists of a glass slide coated with a thin (50 nm) gold film to which is attached, by an inert linker layer, a chemical matrix onto which one of the binding partners can be immobilised. In our experiments, two different chips were used: a sensor chip CM5 and a sensor chip NTA (nitrilotriacetic acid) (figure 68). CM5 sensor chips are the most commonly used chips and have a surface matrix consisting of non-crosslinked carboxy-methylated dextran (Nice and Catimel, 1999). The carboxyl groups are used for covalent immobilisation of the ligands via amine, thiol, aldehyde or carboxyl groups (figure 68).

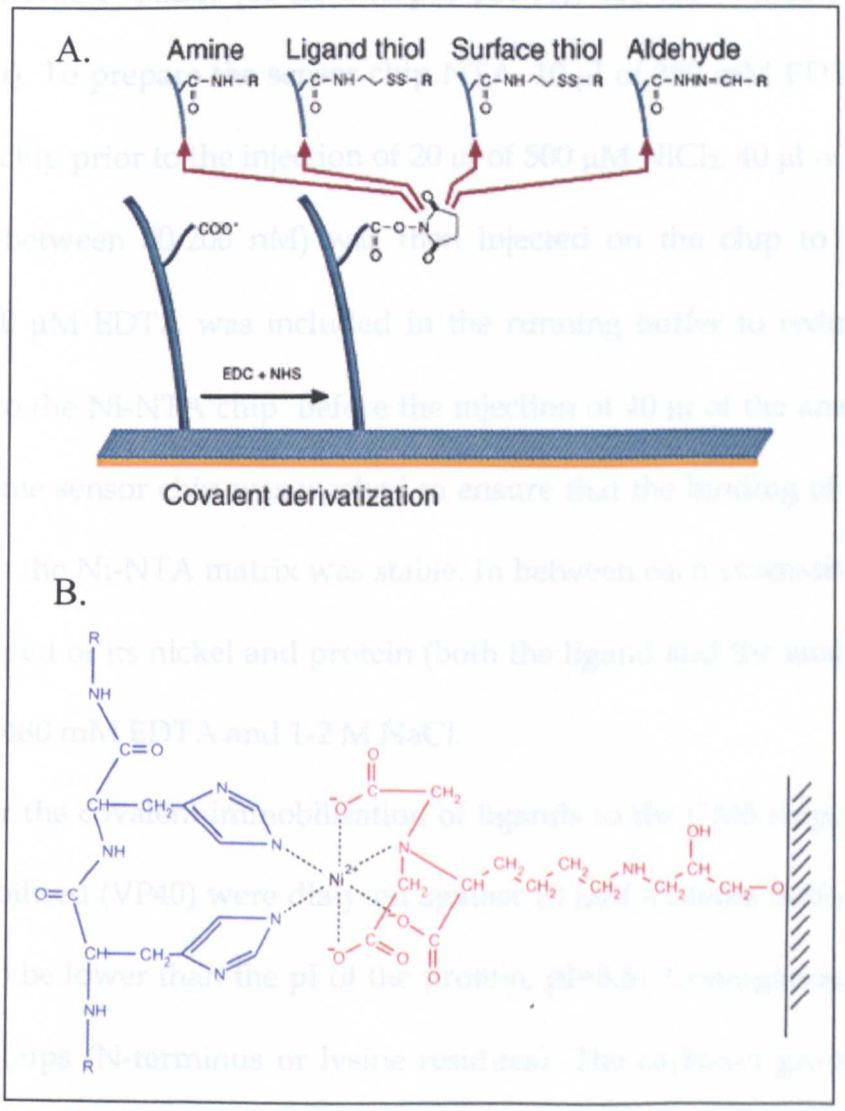

Figure 68: (A) Schematic diagram of the covalent immobilisation of ligands to CM5 sensor chips. Taken from http://www.biacore.com/biomol/pdfs.shtml. **(B)**

Interaction between a His-tagged protein (blue) and the Ni-NTA matrix (red) of a NTA sensor chip.

Sensor chip NTA has a dextran matrix to which NTA is pre-immobilised. The chip is designed to bind His-tagged molecules via chelated nickel for subsequent analysis of analyte binding. Immobilisation via a His-tag has the advantage of orienting the ligand molecules in a homogeneous way.

For our experiments, all the proteins to be analysed were first dialysed against a HBS-P buffer (10 mM Hepes pH 7.4, 150 mM NaCl and 0.005% P20 detergent). To prepare the sensor chip NTA, 10 µl of 350 mM EDTA was injected onto the chip prior to the injection of 20 µl of 500 µM $NiCl_2$. 40 µl of the His-tagged protein (between 50-200 nM) was then injected on the chip to immobilise the ligand. 50 µM EDTA was included in the running buffer to reduce non-specific binding to the Ni-NTA chip. Before the injection of 40 µl of the analyte (Nedd4 or Tsg101), the sensor chip was washed to ensure that the binding of the His-tagged protein to the Ni-NTA matrix was stable. In between each successive run, the chip was stripped of its nickel and protein (both the ligand and the analyte) by washes with 350-480 mM EDTA and 1-2 M NaCl.

For the covalent immobilisation of ligands to the CM5 chip, the proteins to be immobilised (VP40) were dialysed against 10 mM Maleate buffer at pH 6.0 (the pH has to be lower than the pI of the protein; pI=8.5). Conjugation is via primary amino groups (N-terminus or lysine residues). The carboxyl groups of the CM5 matrix were activated by the injection of 50 µl of N-hydroxysuccinimide/ N-ethyl-N'-dimethylaminopropyl-carbodiimide (NHS/EDC), before injection of the ligand (VP40). When a sufficient amount (between 1000 and 3000 RU) of ligand had been

immobilised on the chip, the residual reactive groups were blocked by the injection of 50 µl of 1 M ethanolamine pH 8.5. Finally, the excess or non-covalently linked protein was eliminated by the injection of HBS buffer. The analyte (Nedd4), in HBS-P buffer, was then injected (40 µl) as in the case of the NTA sensor chip. The CM5 chip was regenerated (removal of the analyte only) by the injection of 5 µl of 2 M NaCl, followed by the injection of 5 µl of 0.05% SDS. The same surface could then be used for multiple analyses.

Data are presented as sensorgrams that show the change in resonance units (RU) versus time (s). A signal of 1000 RU is equivalent to a surface concentration of 1 ng/mm^2. A typical sensorgram is shown in figure 69.

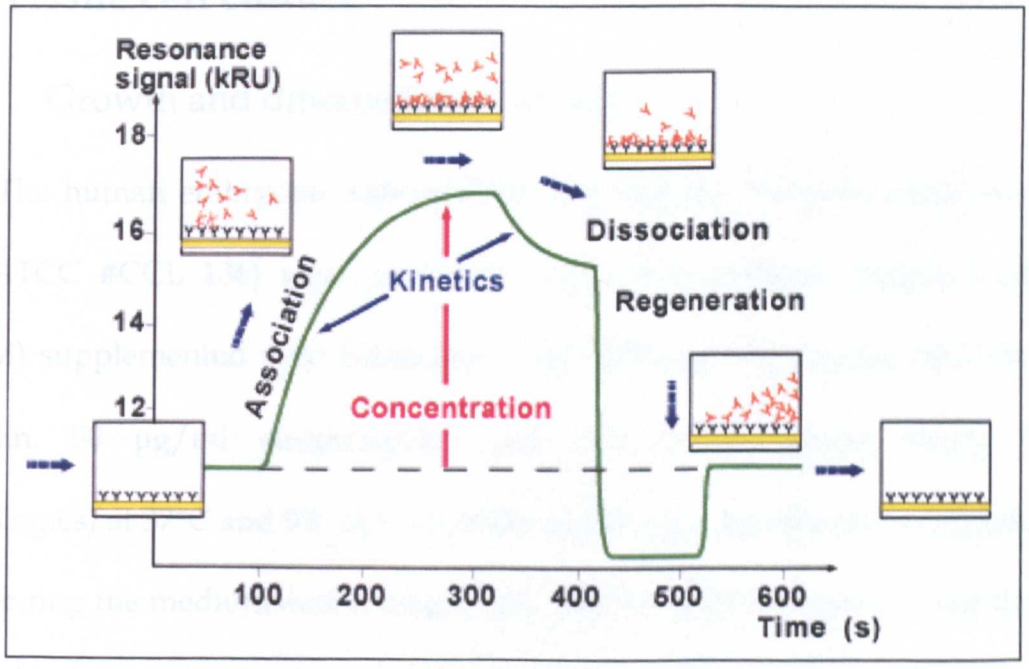

Figure 69: Characteristics of a sensorgram. The sensorgram can be divided into three major phases: association, dissociation and regeneration. Taken from http://www.biacore.com/biomol/pdfs.shtml.

For all the experiments, non-specific binding of the analyte to either the CM5 or NTA sensor chip was tested, and a sensorgram was measured after injection of

HBS-P buffer instead of the analyte. The sensorgrams corresponding to the buffer were then subtracted from those obtained after injection of the analyte. Detailed kinetic analysis requires injecting varying concentrations of analyte over immobilised ligand and analysing the resulting experimental data, using the BIAevaluation software. For the analysis of the sensorgrams, the regeneration phase and the immobilisation of the ligand, in the case of the NTA chip, were removed, and the sensorgrams were brought to zero along the y- and the x-axis, for a proper overlay. The programme then fits the experimental data to a calculated model, in order to obtain the kinetic parameters.

6. Tissue cell culture

6.1 Growth and differentiation of cells

The human embryonic kidney 293T cells and the rhabdomyosarcoma RD cells (ATCC #CCL 136) were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with Glutamax I and 4500 mg/ml glucose, 50 units/ml penicillin, 50 μg/ml streptomycin, and 10% foetal bovine serum (Life Technologies) at 37°C and 5% CO_2. At 100% confluence, the cells were subcultured by decanting the medium and adding trypsin-EDTA (0.25% trypsin, 1 mM EDTA; Life Technologies) for 5 min at room temperature, to detach the cells. The cells were then pelleted by centrifugation at 1100 rpm for 5 min, resuspended in fresh medium and diluted at a 1: 5 ratio.

The RD cells were transferred for several days (up to 5 days) to low serum (1% foetal bovine serum) medium to differentiate. To study their differentiation, the cells were grown on glass coverslips and were treated with the May-Grünwald

and Giemsa stains (Sigma) prior to mounting the slides. This treatment allows a clear visualisation of the nuclei. The extent of differentiation was estimated by counting the number of multinucleated cells (more than 2 nuclei) within a group of 100 cells. Several groups of cells throughout the coverslip were counted.

6.2 Transient and stable transfections

For transient transfections, 293T cells were grown to 50% confluence in 100 mm petri dishes and then transfected with 15 μg of plasmid DNA mixed with 50 μl of Lipofectin (Life Technologies) in serum-free medium (Opti-MEM; Life Technologies). Similarly, RD cells were grown in 6-well plates to 50% confluence and were transfected with 5 μg of plasmid DNA mixed with 10 μl of Lipofectin. For immunofluorescence staining, the cells were seeded 15-20 hours prior to transfection in 6-well plates with glass coverslips. Cells and supernatants were collected three days post-transfection for further analysis.

For stable transfections, the cells were grown to 30-50% confluence in 100 mm plates and were transfected with 1.5 μg of plasmid DNA and 15 μl of Lipofectin diluted in Opti-MEM medium. The following day the DNA-containing medium was replaced with complete DMEM with penicillin and streptomycin. After 48 hours, the cells were subcultured at a 1:5 ratio into selection medium, containing the appropriate antibiotic. The cells were then incubated at 37°C and 5% CO_2 until distinct cell foci could be seen. To isolate these foci, greased cloning cylinders (Polylabo) were placed over the cells. After removing the medium within the cylinders, 20 μl of trypsin-EDTA solution was added to the cylinder to detach the cells. The cells were then resuspended in 100 μl of fresh medium and transferred to 24-well culture plates. This was repeated for several (up to 15) cell

foci. To ensure these were real stable transfectants, the cells were amplified in selection medium and tested for the expression of the gene of interest, either by immunofluorescence or by Western blot.

For the preparation of the ADAM 12/RD stable cell line the protocols described in the T-Rex system (Invitrogen) manual were carefully followed. A first stable cell line was created, by selection with blasticidin, that contained the pcDNA6/TR plasmid. This cell line was then used to select for transfected cells possessing both pcDNA6/TR and pcDNA4/myc-his/ADAM 12, by growing the cells in medium containing both blasticidin and zeocin. Blasticidin and zeocin were added to a final concentration of 7.5 μg/ml and 250 μg/ml respectively. The expression of ADAM 12 was induced by the addition of tetracycline to a final concentration of 1 μg/ml.

6.3 Immunofluorescence staining

Cells were grown on cover glasses and were analysed by indirect immunofluorescence three days post-transfection. Cells were washed three times with cold PBS on ice containing 1 mM $MgCl_2$ and 0.1 mM $CaCl_2$ and subsequently permeabilised and fixed with a solution containing ice-cold methanol (80 %) and acetone (20 %) for 20 min at -20°C. Cells were then blocked with PBS containing 1% BSA for 30 min before being incubated with the primary antibody for one hour at room temperature. Cells were washed three times with cold PBS containing 1 mM $MgCl_2$ and 0.1 mM $CaCl_2$ before adding the secondary antibody conjugated to fluorescein-isothiocyanate (FITC), an anti-rabbit-IgG or anti-mouse-IgG (Life Technologies, Inc.). Cover glasses were mounted in fluoromount-G (Southern

Biotechnology Associates, Inc.) and fluorescence staining was visualised using an Axioskop fluorescence microscope (Zeiss).

6.4 Plasma membrane preparations

The plasma membranes (PM) were isolated using an iso-osmotic homogenisation medium (adapted from (Hubbard et al., 1983)). 72 hours post-transfection the cells were washed and detached in PBS on ice. The pelleted cells were then resuspended in homogenisation buffer (HB), consisting of 0.25 M sucrose, 10 mM Tris pH 7.4 and 1 mM $MgCl_2$. Protease inhibitors (Roche) were also added to the buffer. The cells were homogenised in a Dounce homogeniser. The nuclei were pelleted by centrifugation at 1300 rpm for 10 min. The organelles and plasma membranes were then pelleted at 3800 rpm for 10 min. This pellet was resuspended in HB, before mixing it with 2 volumes of 2 M sucrose in 10 mM Tris pH 7.4 and 1 mM $MgCl_2$. This mix was then overlaid with HB and the membranes were purified by ultracentrifugation in a SW 41 rotor at 24,000 rpm for 1 hour at 4°C. The fraction corresponding to the plasma membranes floats up to the interface. This fraction was collected and diluted 10 times in HB before harvesting the membranes by centrifugation at 8000 rpm for 10 min. The membranes were resuspended in SDS-sample buffer for Western blot analysis.

References:

Brinkmann, U., Mattes, R. E., and Buckel, P. (1989). High-level expression of recombinant genes in Escherichia coli is dependent on the availability of the dnaY gene product. Gene *85*, 109-114.

Dessen, A., Forest, E., Volchkov, V., Dolnik, O., Klenk, H. D., and Weissenhorn, W. (2000). Crystallization and preliminary X-ray analysis of the matrix protein from Ebola virus. Acta Crystallogr D Biol Crystallogr *56*, 758-760.

Hubbard, A. L., Wall, D. A., and Ma, A. (1983). Isolation of rat hepatocyte plasma membranes. I. Presence of the three major domains. J Cell Biol *96*, 217-229.

Karlsson, R., and Falt, A. (1997). Experimental design for kinetic analysis of protein-protein interactions with surface plasmon resonance biosensors. J Immunol Methods *200*, 121-133.

Nice, E. C., and Catimel, B. (1999). Instrumental biosensors: new perspectives for the analysis of biomolecular interactions. Bioessays *21*, 339-352.

Ruigrok, R. W., Schoehn, G., Dessen, A., Forest, E., Volchkov, V., Dolnik, O., Klenk, H. D., and Weissenhorn, W. (2000). Structural characterization and membrane binding properties of the matrix protein VP40 of Ebola virus. J Mol Biol *300*, 103-112.

Sambrook J., F. E. C., Maniatis T. (1989). Molecular cloning.

PAGE/PAGES EXCLUDED UNDER INSTRUCTION FROM THE UNIVERSITY